

EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

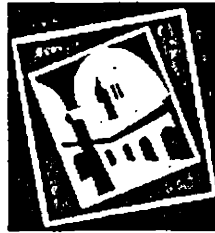
**University Dropouts, Returns to Education,
Job Displacement, and International Risk
Sharing: Four Essays in Empirical
Economics**

Sascha O. Becker

Thesis submitted for assessment with a view to obtaining
the Degree of Doctor of the European University Institute

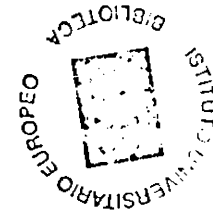
Florence, November 2001

B/c →



EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

LIB
338.4737
011 BEC



**University Dropouts, Returns to Education,
Job Displacement, and International Risk
Sharing: Four Essays in Empirical
Economics**

Sascha O. Becker

The Thesis Committee consists of:

- Prof. Michael Artis, EUI, second advisor
- “ David Card, University of California at Berkeley
- “ Andrea Ichino, EUI, supervisor
- “ Jörn-Steffen Pischke, L.S.E.

Contents

Acknowledgements	vii
Introduction	ix
1 Why Don't Italians Finish University?	1
1.1 Introduction	1
1.2 Some facts	3
1.2.1 Italy	3
1.2.2 Germany	5
1.2.3 The dropout phenomenon in historical perspective	6
1.2.4 Preliminary evidence for Italy based on data about high school graduates	6
1.3 Time-to-educate in a job search model	9
1.3.1 The basic version of the model	9
1.3.2 Introducing different job arrival rates in unemployment and education	13
1.4 Maximum likelihood estimation	16
1.4.1 Maximum Likelihood Estimation for Italy	17
1.4.2 Maximum Likelihood Estimation for Germany	21
1.4.3 Empirical proxies	23
1.5 Results	23
1.5.1 Italy	23
1.5.2 Germany	25
1.6 Conclusion and Policy Implications	26
1.7 Appendix	28
1.7.1 Further empirical exercises	28
1.7.2 Introducing job destruction into the job search model	29
1.7.3 Institutional differences	31
1.7.4 Data description	32

1.7.5	Tables and Figures	35
2	Returns to Education in Germany	51
2.1	Introduction	51
2.2	Theoretical considerations	53
2.3	Educational outcomes and returns to schooling in Germany	54
2.3.1	Some background information using regional data	55
2.3.2	Previous studies	56
2.4	Data description and sample selection	57
2.5	IV Estimation of the Returns to Education	57
2.5.1	Choice of instrument	57
2.5.2	Which Effect Can We Identify? The variable treatment approach to the estimation of returns to schooling	58
2.5.3	IV Estimation Results	60
2.5.4	Internal validity of the instruments	61
2.5.5	External validity of the instruments	64
2.5.6	Characterizing the response function	65
2.6	Summary and conclusions	66
2.7	Appendix	68
3	Displacement in Germany	89
3.1	Introduction	89
3.2	Defining Displacement	91
3.3	The data	92
3.3.1	The German Socioeconomic Panel (GSOEP)	92
3.3.2	Descriptive statistics	93
3.3.3	The outcomes	94
3.4	The econometric approach	95
3.4.1	Treatment of the unemployed	97
3.5	Empirical Analysis	98
3.5.1	OLS results	98
3.5.2	Fixed Effects Regression I	99
3.5.3	Displacement probit	99
3.5.4	Fixed Effects Regression II	100
3.6	Conclusion and outlook	101
3.7	Appendix	104

3.7.1	Algorithm for the estimation of the propensity score	104
3.7.2	Figures and Tables	105
4	Risk-Sharing in the Short Run and In the Long Run	119
4.1	Introduction	119
4.2	A dynamic model of risk sharing	122
4.2.1	PVAR Estimation	124
4.2.2	Permanent shocks and risk sharing	125
4.3	Data and Empirical Implementation	127
4.4	Results	128
4.4.1	Permanent and transitory shocks	130
4.5	Conclusion	132
4.6	Appendix	133
4.6.1	Data Appendix	133
4.6.2	Technical Appendix	134
A	The GSOEP data	143

Acknowledgements

This thesis would not have been possible without the support of many persons. First, I would like to thank my main advisor Prof. Andrea Ichino. His lectures in labor economics got me interested in the subject and he continuously stimulated me in my work. During most of my stay at the EUI, I worked for him as a research assistant and this work helped me in getting to know more and more data sources that I also used for my own work. Prof. Mike Artis was a very accessible co-advisor and I profited a lot from his advice. Both advisors were critical commentators and thereby helped me to continuously improve the thesis.

I spent an important part of the thesis work at the Center for Labor Economics at the University of California at Berkeley. Prof. David Card was extremely helpful and welcomed me as one of *his* students from the very first minute.

Both the EUI labor lunch and David Card's working group provided a stimulating environment in which I was able to report my progress, problems and questions at any time. I want to thank my colleagues at both the EUI and Berkeley for their comments. Many of them have become close friends and we share many more things than just economics. This is particularly true for Inês Cabral-Bauer, Martin Ellison, Winfried Koeniger, Frank Siebern, and Daniela Vuri. Mathias Hoffmann introduced me to economics when I was still studying mathematics and he passed on his enthusiasm for economic questions to me. I would like to thank him for almost 10 years of close friendship.

Two of the papers in this thesis have been written jointly, one of them with Frank Siebern-Thomas and another one with Mathias Hoffmann. Both joint projects turned out to be a pleasant experience despite the fact that my coauthors and I were more than 1000 kilometers apart in both cases when writing the papers.

Seminar participants at the Institute for the Study of Labour (IZA), the Universities of Louvain-la-Neuve, Munich, and Salerno as well as conference participants at the Young Economist Meeting 1999, the IZA Summer School in Labour Economics 1999, the UC Berkeley Poster Conference 2000, the GSOEP Conference 2000, the Annual Conference of the Italian Association of Labor Economists 2000, the ZEW Young Scholars Conference 2001,

the Royal Economic Society Conference 2001, the CEPR European Summer Symposium in Labour Economics 2001 and workshop participants at Statistics Canada have commented on various stages of the work presented here.

I am grateful to my parents who supported me throughout all of my studies. Without the love and continuous support of my wife Claudia many periods of my life as a Ph.D. student would have been much less enjoyable. I dedicate this thesis to her.

Introduction

This thesis contains four papers. While the first three take a micro labor perspective, the fourth takes an international macro perspective. The common theme in the first three papers is the acquisition and destruction of human capital. The fourth was inspired by a seminar given by Jacques Mélitz in the EUI seminar series on European Monetary Integration. Just as the other papers in this thesis, it uses panel data.

The first paper is concerned with explaining the interaction between (university) education and the labour market for youths in Italy and Germany. The second paper looks at the causal effect of education on earnings. The third paper estimates the earnings losses incurred by displaced workers. The fourth paper examines the channels and extent of international risk-sharing in the short run and in the long run.

The first paper is motivated by a striking empirical fact. Italy's university dropout rate of more than 60% is the highest of all OECD countries and thus contrasts sharply with Germany's dropout rate of 25%. I develop a model of university enrollment and job search that helps to understand the differences between the two countries. For Italy, I identify two main groups of dropouts. Misguided students are ill-prepared to obtain an academic degree. Parking lot students drop out as soon as they get the first suitable job offer but obtain a degree in case they never get a job offer throughout their studies. In Germany, only misguided dropouts exist, and there are fewer of them than in Italy. As to the common theme of human capital, this paper shows that in the absence of clearing (youth) labor markets, individuals might over- or underinvest in human capital.

High school completion rates vary considerably across West-German counties (Landkreise) and are highly correlated with measures of schooling infrastructure. In the second paper, we argue that 'place of childhood' as a proxy of schooling infrastructure is a convincing exogenous source of variation in schooling levels that allows us to identify the causal effect of schooling for well-defined subgroups of the population. Using the variable treatment intensity approach exposed by Angrist and Imbens (1995) we find that individuals from 'poor family background' respond most strongly to the instrument 'place of childhood'. Their re-

sponse is further most pronounced at low schooling levels whereas the response of individuals with 'rich family background' is most pronounced at higher schooling levels. Finally, this approach allows us to detect changes in the response function over time.

The literature on human capital distinguishes between general and specific human capital. When being displaced, a worker loses that part of his human capital stock that is not transferable to the new job and incurs a corresponding earnings loss. My third paper adds to the recent contributions about displacement in Germany (e.g. Bender, Dustmann, and Meghir, 1998, and Burda and Mertens, 2001). It exploits the German Socioeconomic Panel in which the type of job separations is reported by workers. I do not find losses caused by displacement for the whole sample. However, this is due to significant heterogeneity between two meaningful subgroups: while workers at high risk of being displaced do not lose by displacement, those at low risk of being displaced lose about 16 percent in earnings in the years following displacement.

The fourth paper uses a panel of 23 industrialised countries to investigate how short-run and long-run income risks are shared and how the source of uncertainty matters for the way this risk gets insured. Surprisingly, short-term and long-term output risks are found to be equally well insured. Transitory shocks get smoothed almost completely whereas permanent shocks remain 80 percent uninsured. We find a somewhat more important role for international capital markets than earlier studies. Whereas our results tie in with some recent theoretical insights and are consistent with empirical findings on home bias in international portfolios, they raise the question why permanent shocks are so hard to insure internationally.

The data sets used in this thesis are described in the relevant chapters. To avoid repetition, an overview of the German Socioeconomic Panel, which I use in three chapters, is given in Appendix A.

To Claudia

Chapter 1

Why Don't Italians Finish University?

1.1 Introduction

After reaching compulsory schooling age, students can either stay in education or drop out. In a world without uncertainty, students would choose their optimal level of schooling once and for all at the beginning of their educational career, taking into account labor market and earnings opportunities, interest rates, personal discount rates, and their ability. In reality, however, many of these factors are uncertain. We therefore observe transitions between education, employment and unemployment in all directions. Education plays an interesting role in so far as it supposedly increases one's labor market opportunities at the same time as it stops him from earning money. This trade-off is well known from the literature on optimal schooling. Yet, if youths do not find a job, education can play a dual role: being a parking lot for people willing to enter the labor market and at the same time of increasing their labor market opportunities. This first effect seems to be particularly true for Italy's university students. Italy's university dropout rate is the highest of all OECD countries. As a consequence Italy also has one of the lowest rates of people holding a tertiary degree. In contrast, Germany's university dropout rate is much lower and overall educational attainment higher than in Italy. In this paper, I address the factors explaining university enrollment and dropout behavior in both countries and present a job search model that takes into account the role of university education.

The paper is related to several strands of literature. Human capital theory provides a theory for the demand for education (Gary S. Becker, 1993). It considers education as an investment good which individuals acquire until the expected returns from an additional year equal the expected costs: in its simplest version students (respectively their parents) choose

the optimal level of education at the beginning of their lives trading off expected direct (tuition fees) and indirect (foregone earnings) costs against expected benefits (higher future earnings) given personal discount rates. In empirical work, often (optimal) schooling choice is not the primary object of study but only used indirectly in studies of returns to education as first stage in instrumental variables earnings regressions where schooling is considered an endogenous variable.

The literature on school enrollment trends (see e.g. Card and Lemieux, 1997 and Card and Lemieux, 2000) explicitly studies educational attainment of different cohorts across regions and time. Card and Lemieux (2000) extend a standard model of optimal schooling choice in two ways. First, they explicitly allow for distaste for schooling, thereby giving more scope to family background variables. Second, they allow for temporary shocks to the (local) labor market that may induce students to drop out earlier or to stay longer than originally planned. My paper is very much related to this literature although it takes an individual-level perspective and a different theoretical approach.

The literature on school-to-work transitions is very much related to this paper. Labor market transition patterns are studied using transition probability matrices, mobility indicators and multinomial logit estimation. Soro-Bonmati (2000) presents results based on these approaches for Germany, Italy and Spain. However, she only considers transitions between 1993 and 1995, thus only one cross-section of transitions, and therefore she cannot control for regional effects.

The paper is organized as follows. In section 2, I present the basic empirical facts. In section 3, I develop a job search model with two skill types, unskilled and skilled, in which the unskilled (high school graduates) can go to university, and become skilled (university graduates). Obtaining a degree and becoming skilled takes time and the model nicely shows how depending on their expected time to completion, some individuals might drop out of education before obtaining a degree if they get a job offer. The model is able to explain transitions between education, employment, and unemployment. Starting from the equations of the theoretical model, in section 4 I describe how the model can be brought to the data using maximum likelihood techniques. In section 5, I present the results of the estimation. In section 6, I conclude and provide some policy implications.

1.2 Some facts

1.2.1 Italy

In Italy, in 1995, the percentage of those aged 25 to 34 holding at least a secondary degree (*diploma* or *qualifica professionale*) was at 59.9% while being equal to 32.3% for the 35 to 65 years old.¹ The percentage of those holding a university entry-certificate (*diplomati*) was at 54.2% for the 25-34 year old and 28.2% for those aged 35 to 65. This shows a trend towards higher education in Italy. Yet, obtaining a tertiary degree is rather unattractive compared to other countries. Non-college tertiary education is basically inexistent. Short degrees equivalent to a bachelor (*laurea breve*) were only introduced recently (1992-93) and reaching a university degree therefore required at least 5 years of study. The low rate of return to education of two to three percent (Flabbi (2000) and Manacorda (2000)) can hardly make up for the foregone earnings incurred during studies. For this reason, the percentage of those holding a university degree was at 12.2% for the 25-34 age group and at 7.7% for the 35-65 year old which is low compared to international standards. There are interesting differences across genders and regions, though. University graduation rates are higher for women than for men. In the south of Italy still nowadays, 10% of all youth do not finish compulsory education. There are also remarkable differences in high school graduation rates and university graduation rates across regions. I will exploit these regional differences in the empirical analysis below.

The overall low university graduation rate contrasts with a very high rate of first-year enrollment (About 70% of all high school leavers go on to university.²) Those two facts put together describe Italy's elevated university dropout rate which is the topic of this study. Most dropouts occur between the first and second year of study. About 25% of all first-year students do not continue in the second year.

Another phenomenon characterizing the difficulty Italian university students face in pursuing their studies is the high percentage of students lagging behind regular study times (*fuori corso*). More than 80% of all graduates need more time than foreseen to finish their degree. In the academic year 1995/96, amongst all graduates, 45.9% finished 3 or more years later than foreseen.³

In figure 1.1, I plot enrollment rates of the 20-24 year old by region and gender over time obtained from the Italian SHIW data. A striking feature of the data is the drop in enrollment

¹This section is based on the author's calculations using the 1995 SHIW data.

²However, this percentage has been slightly decreasing in recent years due to increased tuition fees.

³See "Università e lavoro: statistiche per orientarsi" (Istat, 1999).

in 1989. The first idea that comes to mind is that this might have to do with a boom in the economy that induces a lot of youths to start working instead of studying. This hypothesis is refuted by the data. In figure 1.3, I show GDP in Italy over the 80s and 90s, which does not show any irregularity of this kind. A deeper look into institutional features, however, reveals the true cause of this drop in enrollment in 1989. In the late 80s, the Italian government promoted the so-called *contratti di formazione e lavoro* (CFL), a sort of apprenticeship for youths, and gave huge subsidies to firms to hire young workers. As shown by Adam and Canziani [1998], the number of these special contracts attained its all-time high in the late 80s and later on was reduced because of tighter public budgets.

All of these facts about school and university enrollment have to be seen in the light of the problem of Italian youths to integrate into the labor market. Figure 1.4 shows the development of youth unemployment between 1984 and 1997 for Italy in comparison to Germany. Italy's youth unemployment rate lies consistently above 30% except for the period 1988-1992 in which the *contratti di formazione e lavoro* temporarily absorbed a lot of youths. In contrast, Germany's youth unemployment rate is considerably and consistently lower. Why is youth unemployment so pronounced in Italy?

The Italian youth labor market is characterized by job queues. Prime-age men are highly protected and basically cannot be fired. This prevents competition between youths and adults for existing jobs. Youths can only enter the labour market when some older workers retire or die. It is a weird institutional feature that some labor contracts even arrange for the bequest of an adult's job to his child by renouncing to a fraction of the severance pay.

I will shortly describe some of the institutional features in Italy that have contributed to the creation of job queues for youths. Workers in core industries are highly protected by the so-called *Cassa integrazione* system. During a downturn in the economy, firms can "temporarily" lay off workers. While laid-off workers still remain officially employees of the firm, they receive 80 percent of their previous wage as a kind of unemployment benefit from the state. In the upturn of the economy, the firm has to take these same workers back on their payroll. Until 1991, according to an Italian law, half of any additional hiring had to be allocated from a list of unemployed from the local state employment agency (*Ufficio di Collocamento*). Only the other half was at free choice to the firm.

Another contributing factor to the creation of job queues is the focus of the Italian social assistance on the family and not on the individual. Italian youths cannot claim unemployment benefits unless they have been employed before. Since most of the unemployed youths are first-job seekers, this ties them to their families and at the same time reduces

their search activity because they are "in a safe place".⁴

1.2.2 Germany

In Germany, in 1995, the percentage of those aged 25 to 34 holding at least a secondary degree⁵ was at 64% while being equal to 56.4% for the 35 to 65 year old, based on GSOEP data. Thus, there is a trend towards higher education also in Germany, yet starting from a higher level. In particular, in the 1980s and 1990s, women quickly caught up in educational attainment and in the middle of the 1990s even overtook men in many dimensions: the rate of females starting university is by now slightly higher for women than for men.

Also Germany did not have short degrees (bachelors) until recently (1997).

Germany is well-known for its vocational training system integrating class-room teaching and on-the job training with firms. The higher number of people with an apprenticeship is part of the explanation for the much higher rate of secondary degrees compared with Italy. Looking more specifically at degrees allowing access to university education (*Abitur* or *Fachhochschulreife* in Germany and *laurea* in Italy), the opposite is true: in 1995, 26.2% of 25 to 34 year old Germans had either *Abitur* or *Fachhochschulreife*, while this was true for 17.5% of the 35 to 65 year old.

As for university degrees, Germans again overtake Italians. While 13.4% of 25-34 year old Germans hold a university degree, it is 12.2% of Italians of the same age.

It is important to keep in mind, however, that the German educational system provides for many different educational opportunities. A lot of students, after working for some time, or after doing an apprenticeship, decide to return to formal education. This is known as *zweiter Bildungsweg* (second-chance education). In Italy, this option is much less common: the latest OECD figures (Education at a glance, 2000) show that for the 1998 German cohort, 28% of them will enter university at some point in their life. For Italians, it is 42%. Comparing this to the percentage of people holding a university entry-certificate we see that for Italians the decision to go to university is a now-or-never decision while the German system is much more flexible in allowing many different educational and career paths.

In figure 1.2, I plot enrollment rates of the 20-24 year old by region and gender over time obtained from the German GSOEP data. Again, there is considerable variation in enrollment

⁴See Becker, Bentolila, and Ichino (2001) and Manacorda and Moretti (2001).

⁵i.e. having done at least an apprenticeship (*Lehre*), finished a specialized vocational school (*Berufsfachschule*), or holding an upper secondary schooling degree (*Abitur*) or having the *Fachhochschulreife* allowing access to technical colleges. For more details concerning the different school types see the section describing the school system.

rates across regions, genders and time.

1.2.3 The dropout phenomenon in historical perspective

The Longitudinal Survey of Italian families (*Indagine longitudinale sulle famiglie italiane, ILFI*) is a data set that allows to study the dropout phenomenon in historical perspective. The ILFI questionnaire collects detailed biography information. One question asks for the year of university enrolment, and the following one if a university degree was obtained. One of the possible answers is that the person dropped out of university in which case she is asked the year of dropout. This information allows us to compute dropout rates by year of enrolment several decades back. Figure 1.5 plots dropout rates for five-year periods between 1936-40 and 1986-90. Dropout rates have increased considerably from 15% in the late thirties to 50% in the seventies and eighties. There are many different reasons behind this considerable rise. First of all, with educational attainment being much lower, and in particular the percentage of youths holding an upper secondary degree (making them eligible for university enrolment) lower, earlier cohorts of students were smaller and most likely more selected in a terms of ability (if we think of innate ability as being similarly distributed over time) and therefore less likely to fail. The increase in the percentage of youths holding an upper secondary degree enlarged the pool of students eligible for university. While before 1968, only youths graduating from classical high school were allowed to enter university, after 1968 all upper secondary degrees allowed university entry. Thus, in addition to the general trend in educational attainment, eligibility criteria for university enrolment were relaxed. Second, over much of the post-World War II period, the rise in university dropout rates is accompanied by a similar trend in youth unemployment rates. These two factors, an increase in the number of youths eligible for university and a lack in their labor market prospects immediately after finishing high school (job queues) might have jointly led to the increase in university dropout rates over a long period. While lack of data prevents us from rigorously explaining the historical dimension of the university dropout phenomenon, we can have a closer look at the interaction of labor market conditions and university enrolment and dropout in the late eighties and the nineties, for which we have better data.

1.2.4 Preliminary evidence for Italy based on data about high school graduates

To get a first idea of the forces behind the Italian dropout phenomenon, in this section I am going to exploit a survey data set from the Italian National Statistical Office (Istat). It is

a representative sample of the year 1995 high school graduates. In 1998, a sample of 18843 students was contacted by phone and asked questions about labor market and education experience during high school and in the three years after leaving high school. In addition, their schools were asked to provide information on students' performance in high school. The major drawback of the data is its limitation to just one cohort of *diplomati*. Yet, it is the only existing data source that allows to *directly* study the dropout phenomenon. I give some descriptive statistics in table 1.1.

Of the 8730 high school leavers that enroll in a *corso di laurea* (i.e. a long degree study) between 1995 and 1998, 8092 (i.e. 92.7%) do so immediately after leaving high school, i.e. in fall 1995.⁶ To study the dropout phenomenon I select this subgroup because we can follow it for full three years and, from the information provided in 1998, see who is still in university and who dropped out. So, here we only lose a tiny fraction of students who first did their military service, or worked etc. and only later on went on to university. Using this well-defined subgroup we end up with a sample of 7495 individuals after deleting observations with missing observations on variables of interest.

In table 1.2, the dropout rate is broken up by type of upper secondary school. The differences are striking: the dropout rate amongst students graduating from vocational tracks (*istituto professionale* and *istituto tecnico*) is several times higher than for students graduating from classical high school (*liceo*). Also note that the overall dropout computed from our sample is quite low compared to the OECD figure of more than 60%. This can be explained by several arguments: our dropout rates only apply to dropouts within the first three years of study while the OECD figure applies to the overall dropout rate, i.e. it also includes late dropouts. Additionally, the above OECD figure applies to the peak in dropout rates in 1990 while the present data are for the 2nd half of the 1990s. After a sharp increase in tuition fees in the early 90s first-time university enrollment decreased, most probably lowering the number of misguided students.

In order to identify the factors behind the dropout phenomenon, I estimate a probit model⁷ where the dependent variable is equal to 1 for university dropouts, and 0 otherwise. The results are displayed in table 1.3. The most interesting finding is the significant difference in the probability to drop out across different school types even controlling for performance in both junior and senior high school. Students graduating from vocational secondary tracks (*istituto professionale* or *istituto tecnico*) are much more likely to drop out of university than

⁶This is in sharp contrast with Germany where a lot of students enroll only after an intermediate period of work or apprenticeship.

⁷I also estimated a linear probability model which yielded basically the same results.

are graduates from classical high school (*liceo*). This result raises doubts about the idea of allowing access to university to everyone. Students coming from professional or technical secondary schools on average do not seem to bring enough academic skills to survive in university. Interestingly, not only does the average grade when leaving high school (*voto di maturità*) positively influence the probability to stay in university but also does the average grade in junior high school (*voto di scuola media inferiore*). So, even conditional on the "ability" right before entering university, the "ability" several years before going to university matters. While the number of siblings does not influence the probability to drop out, parents' education - measured by the maximum of mother's and father's years of schooling - matters a lot. The more educated are the parents, the higher the student's probability to stay in university. This result is remarkable because it shows up although we control for performance in both junior and senior high school.⁸ Finally, students that are older at high school graduation have a higher probability of dropping out. Number of classes repeated does not significantly enter once we control for age.

What I take from this exercise is the fact that a lot of students are clearly misguided to university. This is particularly true for students coming from technical or professional secondary tracks, both of which are not supposed to prepare for academic studies in the first place. Answers of dropouts as to why they dropped out (table 1.4) show that a high percentage of students found their studies too difficult and thereby corroborates the finding that too many low ability students are allowed to enter university in the first place. It also highlights the misconception by students about what university is and what their chances are to obtain a degree in the specific subject of their choice.⁹

Yet, it is striking to see that more than one quarter of the dropouts drops out because they already found a job. Additionally, it is interesting that the majority of those who drop out because they found the studies too difficult, are working in 1998, so a large number of them might not *only* have dropped out because the studies were too difficult but also because their job prospects were sufficiently positive.¹⁰ This interpretation is supported by

⁸However, parents' education might pick up the effect of family income, a variable we cannot control for.

⁹For this reason, in 1998, the Italian Ministry of Public Education introduced a system of *pre-registration* (<http://www.istruzione.it/argomenti/orientamento/orpreiscrizioni.htm>) by which students who wish to go to university after leaving high school can informally enroll in a subject of their choice. Then, a close cooperation between high schools and universities is set up: schools are being provided with informative literature and schools organize visits to nearby universities for trial lessons. In Germany, a system of following regular classes at university during the last year at high school to get an idea of what university is like, has been already set up many years ago.

¹⁰This is a standard problem in surveys when only one answer can be given.

the fact that the percentage of students that claim to drop out because they found the studies too difficult is much higher in the North (41.3%) - where labor market conditions are very favorable for youths - than in the South (29.9%) - where the youth labor market problem is most pronounced.¹¹

The evidence collected suggests that there are two main groups of dropouts: *misguided* students, mainly coming from vocational secondary schools, are ill-prepared to obtain an academic degree. *Parking lot* students drop out as soon as they get the first suitable job offer.

In the next section, I present a simple theoretical model capturing the interaction between university education and the labor market that can rationalize why so many high school graduates enroll in university and why such a high fraction of them later drops out.

1.3 Time-to-educate in a job search model

1.3.1 The basic version of the model

Consider a continuous time model with two skill types, unskilled and skilled. We can think of the skilled as holding a university degree and the unskilled as being high school graduates without a university degree. The unskilled can go to education, and obtaining a degree, they become skilled. Unskilled workers can be either unemployed, in education, or employed, while skilled workers can only be unemployed or employed. Denote by U_u the expected present discounted value (PDV) of income of an unskilled unemployed, and by U_s the PDV of income for a skilled unemployed. By W_u and W_s denote the PDV of being an unskilled employed and of being a skilled employed, respectively, and by E_u the PDV of being an unskilled in education. U_u , U_s , W_u , W_s , and E_u can be given asset interpretations and their relationship can be written in the form of arbitrage equations. Remark that I do not model the firm side here. Of course, this could be easily done but I abstract from it for two reasons: for the clarity of the exposition and for the reason that in the empirical part I do not have data on the firm side anyway.

Let b be the value of the outside option (we might think of it as unemployment benefits), w_u the wage rate of the unskilled and w_s the wage rate for the skilled, all of which are taken to be exogenous.¹² By r denote the rate of time preference. Assume that an unskilled

¹¹An alternative explanation would be that universities are more demanding in the North.

¹²Since the number of school-leavers entering the labor market is small compared to the total labor force, we can reasonably consider school-leavers to be price-takers, with w_u and w_s determined by the distribution of skills in the population.

unemployed has a constant probability λ_u of finding a job at any instant, and a skilled unemployed finds a job with instantaneous probability λ_s . Then, we can write the asset equations defining U_u and U_s as

$$rU_u = b + \lambda_u(W_u - U_u) \quad (1.1)$$

$$rU_s = b + \lambda_s(W_s - U_s) \quad (1.2)$$

where for the time being we assume b to be the same for the skilled and the unskilled. A: unskilled can take further education in which case he still receives job offers with probability λ_u that he can accept or reject, and with probability γ_i he obtains a degree and become skilled.¹³ We assume that newly graduated individuals first go into unemployment.¹⁴ Remark that γ is indexed by an individual-specific index i allowing for heterogeneity in "degree achievement rates". This reflects the fact that the expected time for reaching a degree varies considerably by individual. γ_i can be interpreted as individual ability and the setup therefore captures the idea that more able students obtain a degree more quickly than less able students. When the outside option remains b , we can write the asset equation defining E_u as

$$rE_{u,i} = b + \gamma_i(U_s - E_{u,i}) + \lambda_u \max(W_u - E_{u,i}, 0) \quad (1.3)$$

In this first setup, the asset equations for W_u and W_s are very simple:

$$rW_u = w_u \quad (1.4)$$

$$rW_s = w_s \quad (1.5)$$

It is very important to stress that all of the arrival rates λ_u , λ_s , and γ_i in my model are *instantaneous* probabilities. Put differently, they are hazard rates of leaving a certain state, conditional on having been in that state until now. They can take values between 0 and infinity. Also remark that I assume all hazard rates to be constant, i.e. independent of the time elapsed in a given state. This is to say that occurrence of events is regulated by Poisson processes. Poisson processes are known to be memoryless processes. One crucial implication

¹³Implicit in this setup is the assumption that only degrees matter and that *some* education but no degree is no better than no education at all. The assumption that only degrees matter is known as *sheepskin effects*.

¹⁴This assumption is reasonable given that graduates do not start working immediately following their graduation.

of this assumption is that the probability of leaving a state does not increase over time, i.e. one is not more likely to leave a state if he has been in it for a longer time span. For instance, considering the "degree achievement rate" γ_i , this means that conditional on already having been in university for a couple of years one has the same expected time to completion as someone who just enters university (i.e. conditional on having been in university for a very short time). This is of course a very strong assumption which is a better description of enrollment behavior for students at the beginning of their studies. In a future extension to the model, I want to consider other types of processes that allow for increasing hazard rates to make the model more realistic also for later years of study.

To see which economic variables drive the decision to continue education or to drop out, consider an unskilled individual in education. When a job offer arrives, he can either accept or reject it. Given the heterogeneity in "degree achievement rates"/ability, there will be a marginal type of individual who is exactly indifferent between continuing education and dropping out. For this individual, the condition $W_u = E_{u,i}$ holds. Solving equation (1.1) for U_u and equation (1.2) for U_s and substituting in equations (1.4) and (1.5) respectively, we obtain the following expressions

$$U_u = \frac{1}{r + \lambda_u}(b + \lambda_u W_u) = \frac{1}{r + \lambda_u}(b + \frac{\lambda_u}{r} w_u) \quad (1.6)$$

$$U_s = \frac{1}{r + \lambda_s}(b + \lambda_s W_s) = \frac{1}{r + \lambda_s}(b + \frac{\lambda_s}{r} w_s) \quad (1.7)$$

For the marginal individual, the last term in equation (1.3) disappears because of the condition $W_u = E_{u,i}$ and therefore equation (1.3) can be rewritten as

$$E_{u,i} = \frac{1}{r + \gamma_i}(b + \gamma_i U_s) = \frac{1}{r + \gamma_i}(b + \frac{\gamma_i}{r + \lambda_s}[b + \frac{\lambda_s}{r} w_s]) \quad (1.8)$$

The condition $W_u = E_u$ can be expressed as

$$\frac{w_u}{r} = \frac{1}{r + \gamma_i}(b + \gamma_i U_s) = \frac{1}{r + \gamma_i}(b + \frac{\gamma_i}{r + \lambda_s}[b + \frac{\lambda_s}{r} w_s]) \quad (1.9)$$

This expression defines a threshold value γ^d for an individual indifferent between continuing education and dropping out. For individuals with $\gamma_i > \gamma^d$ the last term in (1.3) disappears and they continue education until they obtain a degree. For individuals with $\gamma_i < \gamma^d$ both the second and the third term in (1.3) are "active" and whatever event comes first, degree or job offer, they turn skilled or they drop out.

Even without explicitly solving for γ^d , we can apply the implicit function theorem¹⁵ to see how different parameter values affect individuals at the margin (and thereby also individuals off the margin) in (1.9). We set $b = 0$ for simplicity and can write

$$\frac{\gamma^d}{r + \gamma^d} \frac{\lambda_s}{r + \lambda_s} w_s - w_u = 0 \quad (1.10)$$

Holding all other parameters constant, a marginal increase in the wages of the unskilled, w_u , induces more people to drop out of education:

$$\bullet \gamma^d(w_u) = \frac{(r + \gamma^d)^2 (r + \lambda_s)}{r w_s \lambda_s} > 0.$$

Similarly,

$$\bullet \gamma^d(w_s) = -\frac{(r + \gamma^d) \gamma^d}{r w_s} < 0$$

A marginal increase in the wages of the skilled, w_s , provides an incentive for students to stay in university.

$$\bullet \gamma^d(\lambda_s) = -\frac{(r + \gamma^d) \gamma^d}{(r + \lambda_s) \lambda_s} < 0$$

As the job arrival rate for skilled unemployed, λ_s , goes up, more students tend to continue university. Notice that the job arrival rate for unskilled unemployed, λ_u , does not enter the optimum.

$$\bullet \gamma^d(r) = \frac{\gamma^d (2r + \gamma^d + \lambda_s)}{(r + \lambda_s) r} > 0$$

An increase in the discount rate, r , has a positive effect on staying in university.

Conditioning on the values of all other parameters (which uniquely determine the cutoff level γ^d), differences in the ability distribution $f(\gamma_i)$ will affect the *fraction* of dropouts. If the group of students holding a university-entry certificate is less able in country 1 than in country 2, then we expect more students to drop out of university. This describes the selection issue associated with university entry. If a country allows more students to enter university to begin with (Italy vs. Germany) and we assume innate ability to be the same for both countries, by definition a higher number of less able students enter university than

¹⁵Implicit function theorem: Let $G(x, y)$ be a C^1 function on a ball about (x_0, y_0) in R^2 . Suppose that $G(x_0, y_0) = c$ and consider the expression $G(x, y) = c$. If $(\partial G / \partial y)(x_0, y_0) \neq 0$, then there exists a C^1 function $y = y(x)$ defined on an interval I about the point x_0 such that: (a) $G(x, y(x)) \equiv c$ for all x in I , (b) $y(x_0) = y_0$, and (c) $y'(x_0) = -\frac{(\partial G / \partial x)(x_0, y_0)}{(\partial G / \partial y)(x_0, y_0)}$.

if access was restricted. This in turn explains a lot of the Italian dropout phenomenon. To give an example, assume that γ_i is symmetrically distributed with $E[\gamma_i] = \gamma^d$. As the right tail of the distribution becomes fatter, the fraction of stayers increases while the fraction of dropouts increases in the opposite case. In order to compute the fraction of dropouts from the model, assume that γ_i is distributed over the interval $(0, \infty)$ following a distribution function $F(\gamma_i)$. Then the expected fraction of dropouts is given by $F(\gamma^d)$.

In this first version of the model, only four states out of five actually occur. Unskilled individuals will always prefer education to unemployment because both their outside option b and the job arrival rate λ_u are the same in both states but in education they might also obtain a degree. In reality, however, we *do* observe unskilled in unemployment. This feature of the model can be changed by assuming the job arrival rate in education to be lower than in unemployment. This version of the model is exposed in the following section.

Also, in this first setup, there is no job destruction and therefore all individuals will finally end up in employment (given that wages are taken to be exogenous and therefore will not adjust the increased labor supply). Employment is an absorbing state. Job destruction can be easily introduced in the model by assuming an exogenous job destruction at rate δ_u for unskilled jobs and at rate δ_s for skilled jobs. Equations (1.4) and (1.5) can be modified to read

$$rW_u = w_u + \delta_u(U_u - W_u) \quad (1.11)$$

$$rW_s = w_s + \delta_s(U_s - W_s) \quad (1.12)$$

This modified set of equations (1.1)-(1.3) and (1.11)-(1.12) can again be solved for an expression defining a marginal individual indifferent between staying in and dropping out of education. Details for this setup are shown in the appendix.

1.3.2 Introducing different job arrival rates in unemployment and education

We saw that in the basic version of the model, no unskilled are observed in unemployment because $E_{u,i} > U_u$. When job arrival rates while in education are as big as they are while in unemployment, education is more attractive because it additionally provides for the chance of obtaining a degree.

In equation (1.3) we can introduce a job arrival rate $\eta_u < \lambda_u$. This may be reasonable if being in education is associated with less time for job search than being unemployed. Equation (1.3) then reads

$$rE_{u,i} = b + \gamma_i(U_s - E_{u,i}) + \eta_u \max(W_u - E_{u,i}, 0) \quad (1.13)$$

Now, we want to see when $E_{u,i} \geq U_u$. From the previous analysis (see equation (1.9)) we know that the threshold γ^d is independent from λ_u . We can therefore distinguish two cases

- $\gamma_i < \gamma^d$, and hence $W_u > E_{u,i}$, and
- $\gamma_i > \gamma^d$, and hence $W_u < E_{u,i}$.

In the first case, the last term in (1.13) does not disappear and we can write

$$E_{u,i} = \frac{1}{r + \gamma_i + \eta_u} [b + \gamma_i U_s + \eta_u W_u] = \frac{1}{r + \gamma_i + \eta_u} [b + \gamma_i \{ \frac{1}{r + \lambda_s} (b + \frac{\lambda_s}{r} w_s) \} + \eta_u \frac{w_u}{r}] \quad (1.14)$$

For simplicity set $b = 0$. Then, $E_{u,i} \geq U_u$ reads

$$\frac{\gamma_i}{(r + \gamma_i + \eta_u)} \frac{\lambda_s}{r + \lambda_s} w_s + \frac{\eta_u}{(r + \gamma_i + \eta_u)} w_u \geq \frac{\lambda_u}{r + \lambda_u} w_u \quad (1.15)$$

This equation defines a new threshold value $\gamma^e < \gamma^d$, that determines whether an unskilled prefers to remain unemployed or to go to education. If $\gamma_i < \gamma^e$, the chance of obtaining a degree is so low that it cannot trade off the lower job arrival rate in education. If $\gamma_i > \gamma^e$, the lower job arrival rate in education is set off by a high enough degree achievement rate and therefore makes going to education worthwhile.

The second case is much simpler: since always $U_u < W_u$ but at the same time $W_u < E_{u,i}$ in this second case, we find $U_u < E_{u,i}$ and therefore everyone with $\gamma_i > \gamma^d$ goes to education. This is self-evident after studying the previous case: observing that $\gamma_i > \gamma^d > \gamma^e$ yields the same result.

To sum up: in the case in which the job arrival rate for the unskilled in education is lower than in unemployment, there are three cases:

- $\gamma_i < \gamma^e$: those with a very low ability choose to remain unemployed instead of going to education
- $\gamma^e < \gamma_i < \gamma^d$: in this intermediate case, unskilled individuals choose to go into education but drop out of education as soon as they obtain a job offer

- $\gamma_i > \gamma^d$: unskilled individuals with high ability prefer education to unemployment and stay in education until obtaining a degree even in the presence of job offers

Figure 1.6 describes the possible cases.

Again, we can apply the implicit function theorem to (1.15) to see how different parameter values affect individuals at the margin of enrolling in university or remaining unemployed. Holding all other parameters constant, a marginal increase in the wages of the unskilled, w_u , induces more people to remain unemployed:

$$\bullet \gamma^e(w_u) = -\frac{\frac{r(\eta_u - \lambda_u) - \lambda_u \gamma^e}{r + \lambda_u}}{\frac{\lambda_s w_s - \lambda_u w_u}{r + \lambda_s}} = -\frac{\frac{r(\eta_u - \lambda_u) - \lambda_u \gamma^e}{r + \lambda_u}}{U_s - U_u} > 0.$$

This is the case because the numerator is positive since $\lambda_u > \eta_u$ and the denominator is $U_s - U_u > 0$. In contrast, an increase in the wages of the skilled, w_s , incentivates more students to enroll in education:

$$\bullet \gamma^e(w_s) = -\frac{\frac{\gamma^e \lambda_s}{r - \lambda_s}}{U_s - U_u} < 0.$$

In the same way, as the job arrival rate for skilled unemployed, λ_s , goes up, more students enroll in education:

$$\bullet \gamma^e(\lambda_s) = -\frac{\frac{r - \lambda_s}{r + \lambda_s} \gamma^e w_s}{U_s - U_u} < 0$$

An increase in the job arrival rate for the unskilled unemployed increases the number of people preferring to remain unemployed and the opposite is true for an increase in the job arrival rate while in education:

$$\bullet \gamma^e(\lambda_u) = \frac{r(r + \gamma^e + \eta_u)w_u}{U_s - U_u} > 0$$

$$\bullet \gamma^e(\eta_u) = -\frac{w_u - \frac{\lambda_u}{r + \lambda_u} w_u}{U_s - U_u} < 0$$

In order to compute the fraction of dropouts in this case of the model, assume again that γ_i is distributed over the interval $(0, \infty)$ with distribution function $F(\gamma_i)$. Then the expected fraction of dropouts is given by $\frac{F(\gamma^d) - F(\gamma^e)}{1 - F(\gamma^e)}$.

Remark that in a cross-section of individuals there are two margins affecting enrollment behavior: the unskilled can choose to enroll or not and the enrolled can choose to accept job offers when they arrive or to reject them. The decision to drop out does not change with respect to the basic version of the model. Behavior of individuals at both margins together determines overall enrollment behavior. Interestingly, comparative statics at both margins

give an unambiguous answer on enrollment behavior. For instance, a higher skilled wage both increases the number of individuals who start education ("entry margin") and increases the number of individuals who reject job offers while in education ("exit margin"). Therefore, the unambiguous effect of an increase in skilled wages is a higher fraction of individuals in education. What is ambiguous is the implication for the dropout rate. To see this, consider an increase in the skilled wage w_s : both γ^d and γ^e go down and the shift of γ^d relative to γ^e determines whether $\frac{F(\gamma^d) - F(\gamma^e)}{1 - F(\gamma^e)}$ goes up or down.

In the empirical part, I am going to analyze this version of the model employing a maximum likelihood procedure.

1.4 Maximum likelihood estimation

On the basis of the theoretical model there are four different groups of individuals:

- *never-takers* never enroll
- *actual dropouts* first enroll in university and later on drop out because they obtain a suitable job offer
- *potential dropouts* enroll in university, and are at risk of dropping out but simply do not happen to get a suitable job offer and finally obtain a university degree
- *always-takers* enroll and never drop out, even in the presence of job offers

Every individual belongs to exactly one of the four groups with probabilities $\pi_{i,NT}$, $\pi_{i,AD}$, $\pi_{i,PD}$, $\pi_{i,AT}$ respectively. $Y_{ij} = 1$ if individual i belongs to group j , otherwise it is 0. This means that $\sum \pi_{ij} = 1$ and $\sum Y_{ij} = 1$. The likelihood contribution of individual i is then given by

$$l_i(b; Y_i, X_i) = (\pi_{i,NT})^{Y_{i,NT}} (\pi_{i,AD})^{Y_{i,AD}} (\pi_{i,PD})^{Y_{i,PD}} (\pi_{i,AT})^{Y_{i,AT}} \quad (1.16)$$

To be able to assess the estimation of this empirical model, an ideal data set would contain the following information besides standard personal and family characteristics which I use as proxies for individual ability:

- number and time of job offers received by each individual
- exact wage offered to an individual

Yet, I do not have information on job offers and wage offers by individual. They can only be proxied by some empirical counterpart which I describe in the following section.

Furthermore, ideally we would have a panel data set of high school graduates that contains information for the whole period between their year of graduation from high school until their late twenties. Then, we could study both the enrollment decision after finishing high school and the decision to continue university or to drop out.

In the Italian SHIW data, however, which basically consist of repeated cross-sections¹⁶, we observe every youth only once. The only information we have about her educational career is her highest degree obtained and a dummy variable for being in university or not. This also implies that we are not able to classify her into one of the four groups.

The German SOEP data, in contrast, are a panel data set and we could exploit its longitudinal structure. Still, for reasons of comparability, in this paper I do not employ panel estimators on the German data, either. This will be done in a companion paper (Becker, 2001). What I would like to point out at this stage is that in the absence of information on the exact number and time of job offers, even possessing of panel data does not alleviate the problem of distinguishing *potential dropouts* from *always-takers*. Both groups obtain a degree and the crucial piece of information that would allow us to tell them apart is the missing information on number and time of job offers.

After explaining the main problems in bringing the theoretical model to the data, we can now have a closer look at the actual empirical implementation of the procedure.

1.4.1 Maximum Likelihood Estimation for Italy

In the Italian data, we only observe youths to be in university ($Y_i = 1$) or not ($Y_i = 0$) at one single point in time. Lacking further information about her educational career, a non-enrolled youth could therefore be either a *never-taker* or an *actual dropout*. An enrolled youth could be either an *always-taker* or a *potential dropout* who just did not happen to get a job offer yet. Figure 1.7 illustrates the four cases in the Italian data. Understanding this "problem" also paves the way out. Although we can not tell with certainty which of the four states someone is in, we can attribute the probabilities of being in the respective states. This is equivalent to setting up the likelihood function in the following way:

$$l_i(b; Y_i, X_i) = P(Y_i = 0)^{1-Y_i} P(Y_i = 1)^{Y_i} \quad (1.17)$$

¹⁶From 1989-1995, there is a small component of rotating panel. However, it is too small to be analyzed separately.

where

$$P(Y_i = 0) = P(\gamma_i < \gamma^e) + P(\gamma^e < \gamma_i < \gamma^d) * P(\text{job offer received}) \quad (1.18)$$

and

$$P(Y_i = 1) = P(\gamma_i > \gamma^d) + P(\gamma^e < \gamma_i < \gamma^d) * P(\text{no job offer received}) \quad (1.19)$$

where γ^e is the entry threshold and γ^d is the dropout threshold and γ_i is the "degree achievement rate" (which can be thought of as ability).

In order to implement this likelihood function we make the following assumptions:

Assumption 1 (Age at enrollment)

Students enroll in university immediately after leaving high school or they never enroll, more specifically they enroll at age 19 or never.

This assumption is supported by OECD figures which show that the majority of Italian students enrolls at age 19.

Assumption 2 (External observer information):

- The external observer knows that the thresholds γ^e and γ^d are functions of wages and job arrival rates and of an unobservable part:

$$\gamma^e = \gamma^e(L) * \vartheta^e \quad (1.20)$$

$$\gamma^d = \gamma^d(L) * \vartheta^d \quad (1.21)$$

where L denotes labor market variables and ϑ^e and ϑ^d are distributed according to the cdf function $G(\cdot)$.

- The ability of an individual can be proxied by family characteristics F but there remains an unobservable part as well:

$$\gamma_i = \gamma_i(F) * \vartheta^f \quad (1.22)$$

where ϑ^f is also distributed according to the cdf function $G(\cdot)$.

Assumption 3 (Parametrization of γ^e and γ^d)

In the above analysis we did not provide closed-form solutions for γ^e and γ^d because they are rather involved non-linear functions of the labor market variables. As an approximation to these non-linear relationships and in accordance with the way hazard rates are typically parametrized in duration models, we assume γ^e and γ^d to be loglinear functions of the labor market variables, i.e. $\gamma^e = \exp(L'\varphi)$ and $\gamma^d = \exp(L'\psi)$. To see why this is not an arbitrary parametrization, let us have a look at the Taylor approximation to γ^d . From equation (1.10) it follows that:

$$\gamma^d(a+h) \approx \gamma^d(a) + \frac{\partial \gamma^d}{\partial \omega_u}(a)h_1 + \frac{\partial \gamma^d}{\partial \omega_s}(a)h_2 + \frac{\partial \gamma^d}{\partial r}(a)h_3 + \frac{\partial \gamma^d}{\partial \lambda_s}(a)h_4 \quad (1.23)$$

So we could approximate γ^d by a linear function of the labor market variables. The use of the exponential has the simple reason that we want γ^d to be positive because it represents a hazard rate. This does not change the fact that the signs of the coefficients φ and ψ represent the sign of the derivatives of γ^e and γ^d with respect to the labor market variables.

Assumption 4 (Parametrization of γ_i)

Following the example above, we parametrize $\gamma_i = \exp(F'\kappa)$.

Assumption 5 (Parametrization of the probability of obtaining a job offer while in university)

One part of the likelihood function is the probability that a student receives a job offer before or after t years in university where $t = \text{age} - 19$ following assumption 1. Assuming the hazard rate η_u to be constant, implies an exponential distribution of the spells in university. The spells in university thus follow a cdf $F(t) = 1 - \exp(-\eta_u t)$. We can parametrize η_u in the same way as the other hazard rates above: $\eta_u = \exp(X'\beta)$, where X denotes variables proxying the hazard rate.

Assumption 6 (Distribution of the unobserved components)

To make the model operational, we will assume that

- $G(\cdot)$ follows a lognormal distribution, i.e. ϑ^e , ϑ^d , and ϑ^f can be written as $\exp(\varepsilon^e)$, $\exp(\varepsilon^d)$, and $\exp(\varepsilon^f)$ respectively where ε^e , ε^d , and ε^f are normally distributed with standard deviation σ_e , σ_d , σ_f .
- ε^e , ε^d , and ε^f are independent.

With the above assumptions, the EO is able to identify the probability that the different outcomes will occur as follows:

If $\gamma_i < \gamma^e$, the youth does not enroll in university. The respective probability is

$$P(\gamma_i < \gamma^e) = P(\exp(F'\kappa) \exp(\varepsilon^f) < \exp(L'\varphi) \exp(\varepsilon^e)) \quad (1.24)$$

$$= P(\varepsilon^f - \varepsilon^e < L'\varphi - F'\kappa) = \Phi \left(z_{fe} < \frac{L'\varphi - F'\kappa}{\sqrt{\sigma_f^2 + \sigma_e^2}} \right) = \Phi \left(z_{fe} < L' \frac{\varphi}{\sigma_{fe}} - F' \frac{\kappa}{\sigma_{fe}} \right) \quad (1.25)$$

where we define $\sigma_{fe} \equiv \sqrt{\sigma_f^2 + \sigma_e^2}$ and where $z_{fe} \sim N(0, 1)$.

If $\gamma^e < \gamma_i < \gamma^d$, the youth is in the range of "potential dropouts" and can either be (still) in university ($Y = 1$) or have actually dropped out ($Y = 0$):

The probability of γ_i being in this intermediate range $[\gamma^e, \gamma^d]$ is

$$P(\gamma^e < \gamma_i < \gamma^d) = P(L'\varphi + \varepsilon^e < F'\kappa + \varepsilon^f < L'\Psi + \varepsilon^d) \quad (1.26)$$

$$= P(F'\kappa + \varepsilon^f < L'\Psi + \varepsilon^d) - P(F'\kappa + \varepsilon^f < L'\varphi + \varepsilon^e) \quad (1.27)$$

$$= P(\varepsilon^f - \varepsilon^d < L'\Psi - F'\kappa) - P(\varepsilon^f - \varepsilon^e < L'\varphi - F'\kappa) \quad (1.28)$$

$$= \Phi \left(z_{fd} < \frac{L'\Psi - F'\kappa}{\sqrt{\sigma_f^2 + \sigma_e^2}} \right) - \Phi \left(z_{fe} < \frac{L'\Psi - F'\kappa}{\sqrt{\sigma_f^2 + \sigma_e^2}} \right) \quad (1.29)$$

$$= \Phi \left(z_{fd} < L' \frac{\Psi}{\sigma_{fd}} - F' \frac{\kappa}{\sigma_{fd}} \right) - \Phi \left(z_{fe} < L' \frac{\varphi}{\sigma_{fe}} - F' \frac{\kappa}{\sigma_{fe}} \right) \quad (1.30)$$

where following the previous definitions $\sigma_{fd} \equiv \sqrt{\sigma_f^2 + \sigma_e^2}$ and $z_{fd} \sim N(0, 1)$. The probability of receiving a job offer in the first t years of study is

$$P(\text{job offer received} \leq t) = 1 - \exp(-\exp(X'\beta)t) \quad (1.31)$$

Obviously, the probability of receiving no job offer in the first t years of study is the complementary probability

$$P(\text{no job offer received} \leq t) = \exp(-\exp(X'\beta)t) \quad (1.32)$$

If $\gamma_i > \gamma^d$, the youth will be in university independent of whether he received a job offer or not. The respective probability is

$$P(\gamma_i > \gamma^d) = P(F'\kappa + \varepsilon^f > L'\Psi + \varepsilon^d) \quad (1.33)$$

$$= 1 - P(F'\kappa + \varepsilon^f < L'\Psi + \varepsilon^d) \quad (1.34)$$

$$= 1 - P(\varepsilon^f - \varepsilon^d < L'\Psi - F'\kappa) \quad (1.35)$$

$$= 1 - \Phi \left(z_{fd} < \frac{L'\Psi - F'\kappa}{\sqrt{\sigma_f^2 + \sigma_e^2}} \right) \quad (1.36)$$

$$= 1 - \Phi \left(z_{fd} < L' \frac{\Psi}{\sigma_{fd}} - F' \frac{\kappa}{\sigma_{fd}} \right) \quad (1.37)$$

Assumption 7 (Equal variance assumption)

Assume $\sigma_{fd} = \sigma_{fe}$ which is equivalent to saying $\sigma_d^2 = \sigma_e^2$ because in this case $\sigma_{fd} = \sqrt{\sigma_f^2 + \sigma_d^2} = \sqrt{\sigma_f^2 + \sigma_e^2} = \sigma_{fe}$.

This assumption allows us to identify the coefficients Ψ , κ , and φ up to the scaling factor $1/\sigma_{fd} = 1/\sigma_{fe}$.

Putting together all of these pieces we obtain the likelihood function which is estimated in the sequel.

1.4.2 Maximum Likelihood Estimation for Germany

While the Italian data were essentially repeated cross-sections, the GSOEP data are a panel data set. This has one big advantage and one disadvantage for our purpose. The big advantage is that the GSOEP data are more informative in the sense that we can follow individuals over time and thus distinguish an *actual dropout* from a *never-taker*. Note that in the group of people finishing their degree *potential dropouts* and *always-takers* still cannot be told apart because we do not have the necessary information.

The disadvantage is that in order to make the results as comparable as possible to Italy, we cannot exploit the panel character of the data in the estimation itself. If we would do and use several observations per individual, i.e. employing panel data estimators, we could not be sure whether differences in results across countries are driven by fundamental differences

in economic variables or whether they stem from different estimation techniques. For this reason, I opt for a rather unusual approach. We use the panel dimension only to obtain more information on the exact type of individual to be able to distinguish between *never-takers* and *actual dropouts* and to find out the exact year when students entered university.¹⁷ After doing this, I randomly draw one observation per individual and from then on treat the data as repeated cross-sections.

Now we can set up the likelihood function in a way very similar to the Italian case. The only difference is that we are able to distinguish between three groups in the German data: *never-takers*, *actual dropouts* and the group of people finishing their degree which comprise *potential dropouts* and *always-takers*.

The likelihood contribution of individual i can be written as

$$l_i(b; Y_i, X_i) = (\pi_{i,NT})^{Y_{i,NT}} (\pi_{i,AD})^{Y_{i,AD}} (\pi_{i,FD})^{Y_{i,FD}} \quad (1.38)$$

where the suffix FD denotes youths that finish their degree and where

$$\pi_{i,NT} = P(\gamma_i < \gamma^e) \quad (1.39)$$

$$\pi_{i,AD} = P(\gamma^e < \gamma_i < \gamma^d) P(\text{no job offer received}) \quad (1.40)$$

if we pick an observation for an actual dropout while she is still in university, and

$$\pi_{i,AD} = P(\gamma^e < \gamma_i < \gamma^d) P(\text{job offer received}) \quad (1.41)$$

if we pick an observation for an actual dropout when she has already dropped out,

$$\pi_{i,FD} = P(\gamma_i > \gamma^d) + P(\gamma^e < \gamma_i < \gamma^d) P(\text{no job offer received}) \quad (1.42)$$

and where γ^e is the entry threshold, γ^d is the dropout threshold and γ_i is the "degree achievement rate".

Apart from assumption 1 which is abandoned because we do observe the actual age at enrolment, all other assumptions are adopted from the Italian part.

Now, I can proceed to explain how I constructed the empirical proxies for the labor market variables of the theoretical model.

¹⁷Remember that for Italy we *assumed* that students enter university at age 19 or never, which we proved to be a very reasonable assumption. In Germany, in contrast, the prevalence of apprenticeships makes this an unreasonable assumption. Moreover, we do not have to assume anything about the first year in university if we can easily read it from the data.

1.4.3 Empirical proxies

In the empirical part, I have to choose empirical proxies for the variables of the theoretical model. After a lot of experimentation¹⁸, the following proxies turned out to be relatively good measures. I use the employment-population ratio of 20-24 year old high school graduates as a proxy for the job arrival rate of the unskilled, λ_u , the employment-population ratio of 30-34 year old university graduates as a proxy for the job arrival rate of the skilled, λ_s , and the return to university education (college premium) to measure $\frac{w_s}{w_u}$. All of these labor market variables L vary by region, gender and year. In contrast, the family background variables F vary at the individual level. Parents' education is measured by the years of schooling of the most-educated parent. Log income of the rest of the family (i.e. excluding the income of the youth herself) measures financial resources. Both family background variables are supposed to proxy for the degree achievement rate (or ability), γ_i . Further controls are age and gender.

1.5 Results

Our main interest lies in the variables defining the entry and dropout thresholds and we would like to see whether indeed the thresholds are affected as predicted by the theoretical model. While I do not have specific expectations about the size of the effects, the theoretical model is explicit about the direction of the effects, i.e. gives clear predictions about the sign of the coefficients. Also should the signs of the coefficients be the same at both thresholds. Let us first look at Italy.

1.5.1 Italy

The results (see Table 1.5) are as follows: while at the entry threshold the signs of the coefficients confirm with the predictions of the theory (although not at a high level of statistical significance)¹⁹, at the dropout threshold only the coefficient on the employment-population ratio of high school graduates is in line with the predictions of the model.

At both the entry and the dropout thresholds higher demand for high school graduates seems to be associated with the thresholds moving to the right, i.e. with both less matriculations and more dropouts, which is consistent with the predictions of the theoretical

¹⁸See section 1.7.1 below

¹⁹See section 1.7.1 for a discussion of the shortcomings of these and further empirical proxies that I experimented with.

model. Since the statistical significance is rather low, this result has to be taken with care. At the entry threshold, higher demand for university graduates is associated with more matriculations (although statistically insignificant). The higher the college premium, the more matriculations we observe while at the dropout threshold we observe more dropouts although we would expect the opposite.

The variables characterizing the "degree achievement rate" respectively ability, parents' education and financial resources, are highly significant and point in the right direction, i.e. "good" family background is associated with a higher probability of starting university and with a lower dropout probability.

The job offer rate while in university, η_u , is positively associated with the employment-population ratio of young high-school graduates.

The behavior of labor market variables at the dropout threshold which seem to be at odds with the predictions of the theoretical model, suggests the following interpretation: according to the survey question about motives for dropping out, a considerable share of students drops out of university because they find their studies too difficult. Discouragement of the *misguided* is not implemented into our model which assumes youths to have perfect knowledge of their ability parameter γ_i . Obviously, discouragement results from a misconceived expectation of the time to degree completion. For *misguided* students, it does not even make sense to use university as a (very cheap but still costly) parking lot once they realize that their expected time to completion is very large (and maybe even infinity). The dropout behavior of the *misguided* students is in some sense independent of the labor market situation at the time of dropout. However, to the extent that labor market variables are serially correlated, it is likely that the labor market situation at enrolment which induced those students to enroll in the first place (in accordance with our theoretical model), is still similar to the one that prevails when these *misguided* students drop out (i.e. doing the opposite of *parking lot* students), leading to a "wrongly-signed" impact of labor market variables for this group of dropouts at the dropout threshold. The "correct" behavior of non-*misguided* students might therefore be superimposed by the *misguided* students effect.²⁰ Let us stress that - once we allow for misconception by students - the discouragement effect is a direct consequence of the enrollment behavior of youths who in the absence of job opportunities

²⁰One possibility to check if this explanation is true, is to introduce interaction terms with type of secondary education, so e.g. with a dummy for *liceo*, the type of secondary school which has presumably by far the lowest number of *misguided* dropouts. This exercise, however is only possible for the 1995 (and the recently issued 1998) data, so reducing the variation in local labor market variables because of a shorter panel dimension; this part still remains to be done once I have more disaggregated measures that can make up for the loss in variation due to a shorter time series.

(wrongly) consider university education the more promising alternative to unemployment. Thus, for these students enrollment and dropout are two sides of the same coin. Our estimates for the dropout threshold are therefore not really at odds with the theoretical model but simply complement it and add the missing piece.

To sum up the results for Italy, labor market variables play an important role in explaining university enrollment behavior. Lower demand for ("unskilled") high school graduates as well as higher demand for ("skilled") university graduates increases first-time university enrollment. Higher returns to university education induce more students to enroll. At the dropout threshold, there are most likely two opposed effects: one subgroup of students (rationally) uses university as a parking lot and drops out as soon as they find a suitable job. This group is the one that I am able to describe by our theoretical model. The second subgroup consists of students with wrong expectations about their ability to obtain a university degree. Only after enrolling, they realize their actual ability and then drop out independent of current labor market variables.

All of these results together highlight again that the main problem for Italian youth is the difficulty of finding the first job.

1.5.2 Germany

For Germany, I first tried to estimate the procedure described in section (1.4.2). Depending on the starting values for the parameter vector, the procedure either did not converge at all or led to totally unreasonable parameter values (in the range of 10^4) with huge standard errors (in the range of 10^5). In the case of convergence, different starting values at each case led to very different but equally unreasonable estimates. We had to conclude that the data are simply not informative for our problem. Put differently, the failure to find a well-defined parameter vector maximizing the likelihood function is an indication that the theoretical model underlying it does not describe the German data.

Then I simplified the problem by grouping (1.40) and (1.41) together, i.e. by denoting the likelihood contribution of actual dropouts as

$$\pi_{i,AD} = P(\gamma^e < \gamma_i < \gamma^d) \quad (1.43)$$

So, instead of differentiating between actual dropouts that are still in university and those that already dropped out, I just use the information that they *are* actual dropouts and therefore must be in the medium ability range. When I estimated this simplified version of the model, the problem of non-convergence or convergence to unreasonable parameter values remained the same.

Finally, I further simplified the estimation procedure by "assuming away" potential dropouts, i.e. I wrote

$$\pi_{i,FD} = P(\gamma_i > \gamma^d) \quad (1.44)$$

and thus implicitly assumed all university graduates to be always-takers. Note that by doing these simplifications, I remove all parts of the likelihood which contain the job arrival rate, a key ingredient of our theoretical model! Only now does the maximum likelihood procedure converge and yield estimates of reasonable size (although with the signs not showing any pattern). Yet, what I estimate now is no longer a counterpart of the theoretical model and the estimates can basically not be interpreted at all.

But this is exactly what we should expect from a priori reasoning about the German case. The theoretical model was developed to yield a description of the Italian dropout phenomenon, and the results for Germany therefore do not come as a surprise. In Germany, youths do not have problems of integration into the labor market to the extent of Italian youths. As I pointed out above, a very common career path for German high school graduates is to first do an apprenticeship to acquire a vocational certificate which also gives them a backup position when going to university. University enrolment in general seems to be much more driven by a deep interest in the subject and by long-run perspectives as a university graduate than by short- or medium-run fluctuations of labor demand and wages as is the case in Italy.

1.6 Conclusion and Policy Implications

The motivation of this paper was to explain the striking difference in university dropout rates between Italy and Germany. I presented a model that helped to understand the interactions between job search and university education. It highlighted the economic and institutional mechanisms inducing many high school graduates to first enroll in university and later drop out. The model allowed me to identify two main groups of dropouts in Italy. *Misguided* students are ill-prepared to obtain an academic degree. Their decision to enroll in university follows from the impossibility of finding a job and can only be explained by their misconception of their own ability. Their dropout decision is the consequence of the hopelessness to obtain a degree and therefore independent of the labor market situation at the time of dropping out. *Parking lot* students in contrast drop out as soon as they get the first suitable job offer but obtain a degree in case they never get a job offer throughout their studies. In Germany, only misguided dropouts exist, and there are fewer of them than in

Italy. There are no *parking lot* students in Germany because German high school graduates do not have a problem in entering the labor market, anyway.

As for policy conclusions, it seems that the Italian dropout problem has to be approached from many different angles at the same time. First, *allowing* access to university to graduates from basically all secondary school tracks contributes to the high number of youths who *do* enroll. This increases the pool of potential dropouts and leads to a high number of misguided youths in university who are ill-prepared for academic studies and after some time give up. While they are enrolled, they crowd university and deteriorate the study conditions of all other students. In contrast, in Germany, not all secondary school tracks lead to university and therefore, the pool of youths enrolling in university can be expected to be more able on average. In addition, for many over-crowded courses of studies, in Germany access is restricted by means of the so-called *numerus clausus*, the minimum requirement for one's high school grade point average. In Italy, the *numerus clausus* also exists but is not as common as in Germany.

Second, Italy should provide for a vocational training system similar to the German one which equips youths with the skills appreciated in the labor market and thereby fills the gap between *only* having a secondary schooling degree and having a university degree. This could be the key to improving labor market entry for Italian youth and lowering the number of *parking lot* students.

Third, the system of pre-registration that I mentioned in section (1.2.4) has to be continued and even extended. Since it was only introduced in 1998, it is too early to see its effects on enrollment and dropout behavior, but it is very likely to decrease the number of *misguided* students.

1.7 Appendix

1.7.1 Further empirical exercises

I spent several months experimenting with different sample selection criteria and tried many different proxies for the labor market variables L . All of them have shortcomings in one or the other way and eventually did not improve the results presented above. Yet, I want to document the work done because it consumed a major part of the time spent on this thesis chapter.

There might be several reasons why the above results are not fully satisfactory. First, the fact that we cannot tell apart *misguided* from *parking lot* students does not allow us to exclude the *misguided* students from the analysis, thereby leading to the odd results at the dropout threshold (see the discussion in the results section above). One possibility to discard a great deal of the *misguided* students from the sample would be to concentrate on students coming from *liceo*, for whom the main motive to enroll and drop out can be supposed to be labor market conditions and not their misconception of their abilities or their inability to succeed in university (also see table 1.2). Unfortunately, the SHIW only contains information about the type of high school degree obtained in 1995 and 1998. Restricting the analysis to these years, reduced the sample size and time series variation in the labor market variables a lot and results thereby did not improve.

Second, and more generally, the empirical proxies as such might be unsatisfactory. Employment population rates are stock variables while job *arrival* rates are flow variables. One alternative therefore is to use employment-*growth* rates. Since the SHIW data are only biennial, employment-*growth* rates are computed over a span of two years. This again did not improve the results. To see if the crucial problem lies in the two-year span that might not well enough proxy for job opportunities in the year of observation, I tried to compute employment-*growth* rates from another data set, the *Indagine sui consumi delle famiglie*. This is a yearly data set on consumption expenditures provided by ISTAT that also contains information on employment, education and has geographical indicators for the 5 aggregated areas. The results did not improve. Employment-*growth* rates measure the difference between job creation and job destruction. For first-time job-seekers, job creation rates are probably the more relevant measure of job opportunities, i.e. job arrival rates.

It was very difficult to find data on job creation rates at all. The only data on job creation for the whole of Italy and covering all sectors are provided by the Ministry of Employment and Social Affairs. The data are on the number of people starting work mediated by the employment agencies. Unfortunately, the data are not split up by educational degrees but

by 4 broad occupation levels, apprentices (*apprendisti*), non-qualified workers (*operai non qualificati*), qualified workers (*operai qualificati*), and employees (*impiegati*). These groups do not match with our educational levels of interest. Taking employees for university graduates and qualified workers for high school graduates is certainly not the right thing to do and the estimates stemming from this exercise are again unsatisfactory.

To sum up, it is pretty difficult to find empirical proxies for job arrival rates of high school graduates and university graduates searching for the first job, by gender and for their labor market of reference. The empirical results presented in section 1.5 can therefore only be seen as a first attempt to bring the theoretical model to the data. Better data is needed to get more satisfactory results.

1.7.2 Introducing job destruction into the job search model

In section (1.3.1) I proposed an extension of the model to incorporate job destruction built on the modified set of equations (1.1)-(1.3) and (1.11)-(1.12). Subtracting (1.1) from (1.11) and (1.2) from (1.12) and rearranging we obtain

$$(W_u - U_u) = \frac{(w_u - b)}{r + \lambda_u + \delta_u} \quad (1.45)$$

$$(W_s - U_s) = \frac{(w_s - b)}{r + \lambda_s + \delta_s} \quad (1.46)$$

which we can substitute back into equations (1.1) and (1.2) to obtain expressions for U_u and U_s :

$$U_u = \frac{1}{r} \left[b + \lambda_u \frac{(w_u - b)}{r + \lambda_u + \delta_u} \right] \quad (1.47)$$

$$U_s = \frac{1}{r} \left[b + \lambda_s \frac{(w_s - b)}{r + \lambda_s + \delta_s} \right] \quad (1.48)$$

Again, the marginal individual is indifferent between continuing education and dropping out. Thus again, we compare

$$W_u = \frac{1}{r} \left[w_u + \delta_u \frac{(b - w_u)}{r + \lambda_u + \delta_u} \right] \quad (1.49)$$

to

$$E_{u,i} = \frac{1}{r + \gamma_i} [b + \gamma_i U_s] = \frac{1}{r + \gamma_i} \left[b + \gamma_i \left\{ \frac{1}{r} \left[b + \lambda_s \frac{(w_s - b)}{r + \lambda_s + \delta_s} \right] \right\} \right] \quad (1.50)$$

The equation

$$\frac{1}{r} \left[w_u + \delta_u \frac{(b - w_u)}{r + \lambda_u + \delta_u} \right] = \frac{1}{r + \gamma_i} \left[b + \gamma_i \left\{ \frac{1}{r} \left[b + \lambda_s \frac{(w_s - b)}{r + \lambda_s + \delta_s} \right] \right\} \right] \quad (1.51)$$

defines a threshold value γ^d for an individual indifferent between continuing education and dropping out. Individuals with $\gamma_i > \gamma^d$ will continue education while individuals with $\gamma_i < \gamma^d$ will accept the first job offer they get or finish their studies whichever event comes first.

We simplify by setting $b = 0$:

$$w_u \left[1 - \frac{\delta_u}{r + \lambda_u + \delta_u} \right] = \frac{\gamma_i}{r + \gamma_i} \frac{\lambda_s w_s}{r + \lambda_s + \delta_s} \quad (1.52)$$

Applying the implicit function theorem to (1.52) we can see how different parameter values affect individuals at the margin of dropping out of education:

$$\bullet \gamma^d(w_u) = - \frac{1 - \frac{\delta_u}{r + \lambda_u + \delta_u}}{\frac{r}{(r + \gamma^d)^2} \frac{\lambda_s w_s}{r + \lambda_s + \delta_s}} = \frac{1 - \frac{\delta_u}{r + \lambda_u + \delta_u}}{\frac{r}{(r + \gamma^d)^2} \frac{\lambda_s w_s}{r + \lambda_s + \delta_s}} > 0$$

The higher w_u , the more dropouts there are.

$$\bullet \gamma^d(w_s) = \frac{-\frac{\gamma^d}{r + \gamma^d} \frac{\lambda_s}{r + \lambda_s + \delta_s}}{\frac{r}{(r + \gamma^d)^2} \frac{\lambda_s w_s}{r + \lambda_s + \delta_s}} = - \frac{(r + \gamma^d) \gamma^d}{r w_s} < 0$$

Higher skilled wages w_s are an incentive to continue education.

In contrast to the basic version of the model, here also λ_u plays a role because of the possibility of being fired on an unskilled job and the subsequent option to re-enter university.

$$\bullet \gamma^d(\lambda_u) = \frac{\frac{\delta_u w_u}{(r + \lambda_u + \delta_u)^2}}{\frac{r}{(r + \gamma^d)^2} \frac{\lambda_s w_s}{r + \lambda_s + \delta_s}} > 0$$

The higher the probability of finding a job as an unskilled unemployed, the more dropouts there are.

$$\bullet \gamma^d(\lambda_s) = \frac{-\frac{\gamma^d}{r + \gamma^d} \frac{(r + \delta_s) w_s}{r + \lambda_s + \delta_s}}{\frac{r}{(r + \gamma^d)^2} \frac{\lambda_s w_s}{r + \lambda_s + \delta_s}} = - \frac{(r + \gamma^d) \gamma^d (r + \delta_s)}{r \lambda_s} < 0$$

As in the basic case, the higher the probability of finding a job after obtaining a university degree, the less dropouts there are.

$$\bullet \gamma^d(\delta_u) = \frac{-\frac{(r + \lambda_u) w_u}{(r + \lambda_u + \delta_u)^2}}{\frac{r}{(r + \gamma^d)^2} \frac{\lambda_s w_s}{r + \lambda_s + \delta_s}} < 0$$

The higher the job destruction rate for unskilled jobs, the less dropouts there are. The opposite is the case for the job destruction rate of skilled jobs:

$$\bullet \gamma^d(\delta_s) = \frac{\frac{\gamma^d}{r+\gamma^d} \frac{\lambda_s w_s}{r+\lambda_s+\delta_s}}{\frac{r}{(r+\gamma^d)^2} \frac{\lambda_s w_s}{r+\lambda_s+\delta_s}} = \frac{(r+\gamma^d)\gamma^d}{r} > 0$$

As in the basic case above, we can compute the expected fraction of dropouts assuming that γ_i is distributed over the interval $(0, \infty)$ with distribution function $F(\gamma_i)$. Again, the expected fraction of dropouts is given by $F(\gamma^d)$.

1.7.3 Institutional differences

In this section, I describe the institutional differences in the Italian and German school systems. The most remarkable difference is in compulsory schooling age. While in Italy it is at age 14, in Germany it is at age 18.²¹

The Italian school system

Italian children first attend *scuola materna* from the age of 3 to age 6 (equivalent to kindergarten). Then, they move on to the *scuola elementare* for a period of 5 years. After this there is the *scuola media inferiore*, which lasts another 3 years. Its completion marks the end of compulsory education. Students who want to continue their education then go to the *scuola media superiore*, which consists of a number of different establishments offering different specializations: these are the *liceo classico*, the *liceo scientifico*, *liceo linguistico*, and the *istituto tecnico*, which take a further five years, the *scuola magistrale* and *istituto magistrale*, which takes 3 and 4 years, and the *istituto professionale* which takes between 3 and 5 years. The *istituto d'arte* and the *liceo artistico* are special arts schools preparing for the Academy of Arts. All of the tracks with less than 5 years can be "upgraded" by taking one or two additional years of schooling to complete a full 5 years of upper secondary education which is the only prerequisite for entering university.

As for vocational training, there are two different types of possibilities, both of which do not yield formal certificates. First, the classical *apprendistato* (apprenticeship) constitutes the start into a blue-collar career. It gives youths a contract of undetermined length. To firms, it is rather unattractive because youths immediately have all rights of a blue-collar

²¹The regulations differ slightly across German *Länder* (regions) but in all of them full-time education has to be taken until age 16 and part-time education until age 18. Instead of taking part-time education between age 16 and 18 one can also take full-time schooling until age 17.

worker. On the other hand, the *contratti di formazione e lavoro* are contracts of predetermined length that are subsidized by the state.

For more details see Table 1.6.

The German school system

After 4 years of elementary school, students can choose between three main tracks of secondary schooling, *Hauptschule*, *Realschule*, and *Gymnasium*, taking 5, 6, and 9 years respectively. Successfully graduating from *Gymnasium* students obtain the *Abitur*, the high school diploma which gives access to university.

Following *Hauptschule* or *Realschule*, but also following *Gymnasium*, students can start an apprenticeship which combines on-the-job training with class-room education. Apart from this classical German type of vocational training, there are several specialized vocational schools like health care schools but also schools providing training for civil servants.

At the university level, until 1997, short degrees (bachelors) did not exist.

A more detailed picture of the German school system is given in Table 1.7.

1.7.4 Data description

In addition to the Istat data on Italian high school graduates that I extensively discuss in section 1.2.4, I use two other data sets: the Bank of Italy *Survey of Household Income and Wealth* (SHIW) and the *German Socioeconomic Panel* (GSOEP).

The GSOEP data

In this paper, I only use the West-German and foreign subsample because of problems in comparability of the West- and East-German school systems. Also is the time series dimension for these subsamples longer than for the East German and immigrant samples. Using the years 1985-1995, the GSOEP data cover approximately the same years as the Italian SHIW data.

The GSOEP data also only provide a direct question on parents' education in one cross-section, in 1986. However, we can update this information in later years because we can follow individuals over time. This is an advantage over the Italian data. Yet, also in Germany 78% of youths aged 16-24 live at home.

Information about income of the rest of the family can be constructed similarly to the Italian data.

The regions used are nearly identical to the West-German states, the so-called *Länder*. Only the smallest *Länder*, i.e. the cities of Hamburg and Bremen and the Saarland had to be merged with neighboring states because their sample sizes would have been too small to obtain reliable estimates. I end up using the following 8 regions: Berlin, Schleswig-Holstein/Hansestadt Hamburg, Niedersachsen/Hansestadt Bremen, Nordrhein-Westfalen, Hessen, Rheinland-Pfalz/Saarland, Baden-Württemberg, Bayern.

The SHIW data

The survey is available from 1977 onwards and has been run on a yearly basis until 1987 (with the exception of 1985) and every other year since then. Although starting in 1989 a small component of rotating panel is introduced in the survey, I ignore this feature of the data here.

The SHIW data have several shortcomings that partly determine the selection of the sample and also the econometric approach.

First, before 1989, the survey is focusing on income-recipients and therefore only a very reduced set of variables is provided for non-income recipients. For instance, the current schooling degree held is only known for income recipients, i.e. basically for the working population. Since by definition, in the SHIW data one can not declare to be working *and* to be a student at the same time, for persons that declare to be students we do not know their current highest degree. So, for example, for a 19-year old, we can not say whether he is a student still in high school (holding a junior high school degree) or already in university (holding a high school degree). So, in order to be able to select the subsample of youth holding a high school degree, we have to restrict the analysis to the years 1989 and after.

Second, family background variables are not immediately at hand. In the SHIW data, a direct question about parents' education and occupation was only asked in 1995. To be able to use data for more years than just the 1995 cross-section, I therefore had to restrict the analysis to youths living at home. Doing this, I am able to use information provided by the parents themselves. This should not influence the results too much considering that in Italy, 88% of youths aged 16-24 are still living at home.

Possible family background variables are parents' education and occupation and family income. Here, I use the maximum of mother's and father's years of education and family income of the rest of the family, i.e. excluding the youths' labor income, because non-enrolled youths are more likely to earn labor income than enrolled youths, thereby introducing a bias in the estimation. I express family income in 1998 Lira and then transform it into Euro. Doing the same for Germany makes the results directly comparable.

As regions I use five aggregated areas *North-West*, comprising Piemonte, Val d'Aosta, Liguria, and Lombardy, *North-East*, comprising Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, and Emilia Romagna, *Center*, comprising Tuscany, Umbria, Marche, and Lazio, *South*, comprising Campania, Abruzzo, Molise, Puglia, Basilicata, and Calabria, and *Islands*, comprising Sardegna and Sicily.

1.7.5 Tables and Figures

Table 1.1: Some descriptive statistics for the sample of high school graduates

Variable	Mean	Std. Dev.
Age (in years) in 1995	19.73	2.47
Female	0.55	0.50
Live at home in 1998	0.88	0.32
Married in 1998	0.04	0.20

Source: Istat data on high school graduates of the year 1995.

Table 1.2: Dropout rates by type of secondary school

School type	Dropout rate	N
Istituto professionale	43.57	482
Istituto tecnico	28.36	2447
Liceo	8.06	3944
Other	17.04	622
Total	17.72	7495

Notes: fraction of students who enrolled in university in 1995 and dropped out within 3 years. Source: Istat data on high school graduates of the year 1995.

Table 1.3: The probability of dropping out

	marginal effect
Istituto professionale	0.308 (0.028)**
Istituto tecnico	0.180 (0.012)**
Other school	0.0876 (0.021)**
GPA (senior) high school	-0.008 (0.000)**
GPA junior high school	-0.025 (0.005)**
Number of siblings	-0.002 (0.005)
Parents' education	-0.007 (0.001)**
Age	0.024 (0.006)**
Number of classes repeated	0.005 (0.008)
Female	-0.008 (0.015)
Number of observations	7495
R^2	0.16

Notes: Probit estimation. Dependent variable: 1=dropped out of university between 1995 and 1998, 0 if not. Reference school type is liceo. Standard errors in parentheses. * significant at 5% level; ** significant at 1% level. Source: Istat data on high school graduates of the year 1995.

Table 1.4: Motive for dropping out of university

Motive for dropping out	Frequency	Percent
Studies too difficult / didn't like	475	35.77
To do different studies	62	4.67
Studies didn't promise professional opportunities	37	2.79
Studying was too costly	67	5.05
Work	343	25.83
Personal Motives	184	13.86
Military Service	106	7.89
Other	54	4.07
Total	1328	100

Source: Istat data on high school graduates of the year 1995.

Table 1.5:
Results from Maximum Likelihood Estimation

Never-takers	Entry threshold	(Potential) dropouts	Dropout threshold	Stayers
$\gamma_i < \gamma^e$	γ^e	$\gamma^e < \gamma_i < \gamma^d$	γ^d	$\gamma_i > \gamma^d$
Employment- population ratio HS Grads	0.654 (0.354)		0.678 (0.535)	
Employment- population ratio Uni Grads	-0.319 (0.245)		0.284 (0.295)	
Returns to University Education	-0.532 (0.414)		0.579 (0.422)	
Sex	-0.222 (0.087)		-0.161 (0.141)	
<hr/>				
Parents' education	-0.091 (0.005)		-0.091 (0.005)	
Rest of family income	-0.230 (0.038)		-0.230 (0.038)	
Constant	3.318 (0.450)		3.318 (0.450)	

Job offer rate

Empl.-pop. Ratio HS Grads	3.065 (1.197)
Sex	0.434 (0.384)
Constant	-2.550 (0.726)

Note: Maximum likelihood estimation of the "Ordered Probit" model described in section 4. The dependent variable is an indicator taking value 1 if a youth is enrolled in university and 0 if he is not enrolled.

Table 1.6: The Italian Educational System

English Term (Diploma in Brackets)	Usual Years of Schooling, Training	Italian Term (Diploma in Brackets)
primary school	5	scuola elementare (licenza elementare)
lower secondary school (leaving certificate)	8	scuola media inferiore (licenza media inferiore)
upper secondary specialized school	13	istituto tecnico (diploma istituto tecnico)
academic secondary school / "high school" (university entry-level leaving certificate)	13	liceo classico, scientifico e linguistico (maturità)
upper secondary school for arts (requires licenza media inferiore)	+3 or +4	istituto d'arte, liceo artistico (diploma artistico)
vocational school (requires licenza media inferiore)	+3	istituto professionale (diploma professionale)
school for the formation of primary school teachers	+3	istituto magistrale (diploma magistrale)
additional preparation year(s) for university following vocational school, arts schools or istituto magistrale [requires diploma of the mentioned schools]	+1 or +2	anno integrativo (maturità)
health care school	+3	scuola infermieri (diploma infermieri)
university (BA, MA)	+4 or +5 (depending on subject)	università (laurea)
doctorate (Ph.D.)	+3	dottorato
apprenticeship (NO formal certificate!)	one or more years	apprendistato (NO formal certificate!)
special contract for youth of predetermined length with subsidies to the firm	one or more years	contratti di formazione e lavoro (NO formal certificate!)
civil servant school	+1 or +2	

Table 1.7: The German Educational System

English Term (Diploma in Brackets)	Usual Years of Schooling, Training	German Term (Diploma in Brackets)
primary school	4	Grundschule
lower secondary school (leaving certificate)	9	Hauptschule(Hauptschulabschluss) Volksschule (Volksschulabschluss)
intermediate secondary school (leaving certificate)	10	Realschule (Mittlere Reife)
comprehensive school / non-streamed secondary school	9-13	Gesamtschule
upper secondary specialized school (certificate of aptitude for specialized short course higher education)	12	Fachschule (Fachhochschulreife)
academic secondary school / "high school" (university entry-level leaving certificate)	13	Gymnasium (Abitur)
part-time vocational school	+2	Berufsschule
technical college / commercial college (vocational extension certificate)	+2	Fachschule, Handelsschule (Fachschulreife)
specialized vocational school [requires lower or intermediate high school diploma]	+2	Berufsfachschule
health care school	+2	Schule des Gesundheitswesens
civil servant school	+1.5	Beamtenausbildung
polytechnical	+3	FH Ingenieurschule
university (BA, MA, PhD)	+5	Universität, Hochschule (Diplom, Magister, Doktor)
comprehensive university, specialized college	+5	Gesamthochschule
apprenticeship, agricultural apprenticeship	+1.5	gewerbliche, landwirtschaftliche Lehre
trade/commerce apprenticeship	+1.5	kaufmännische Lehre

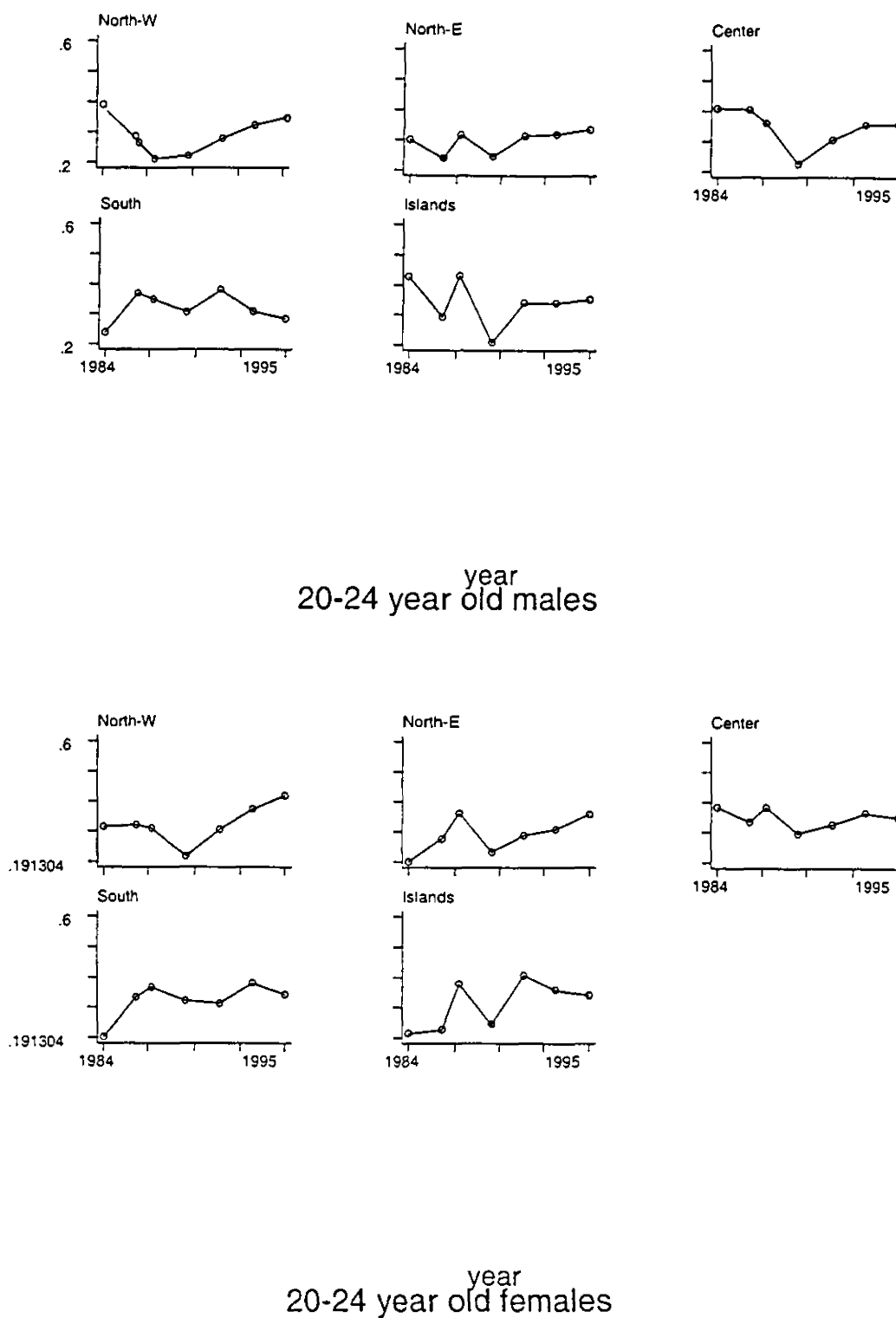


Figure 1.1: enrolment of Italian youth (20-24 year old)

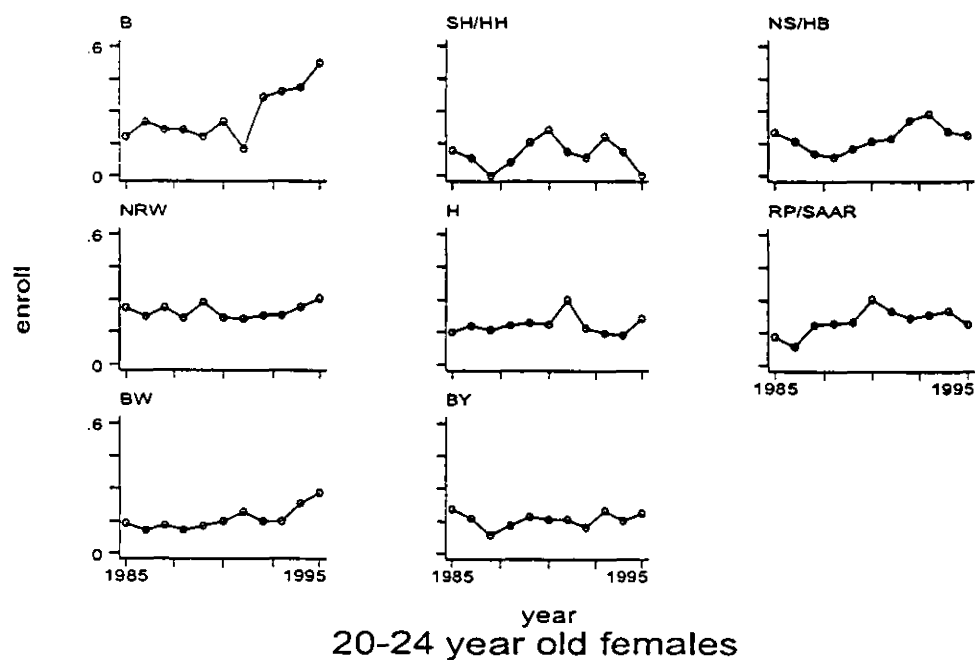
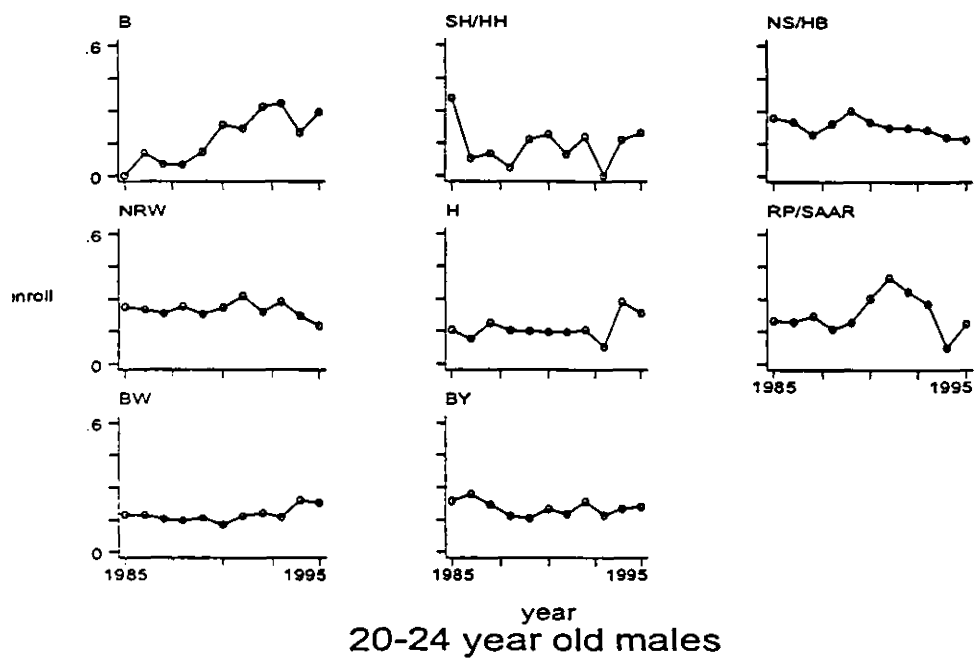


Figure 1.2: enrolment of German youth (20-24 year old)

Figure 1.3: GDP Italy

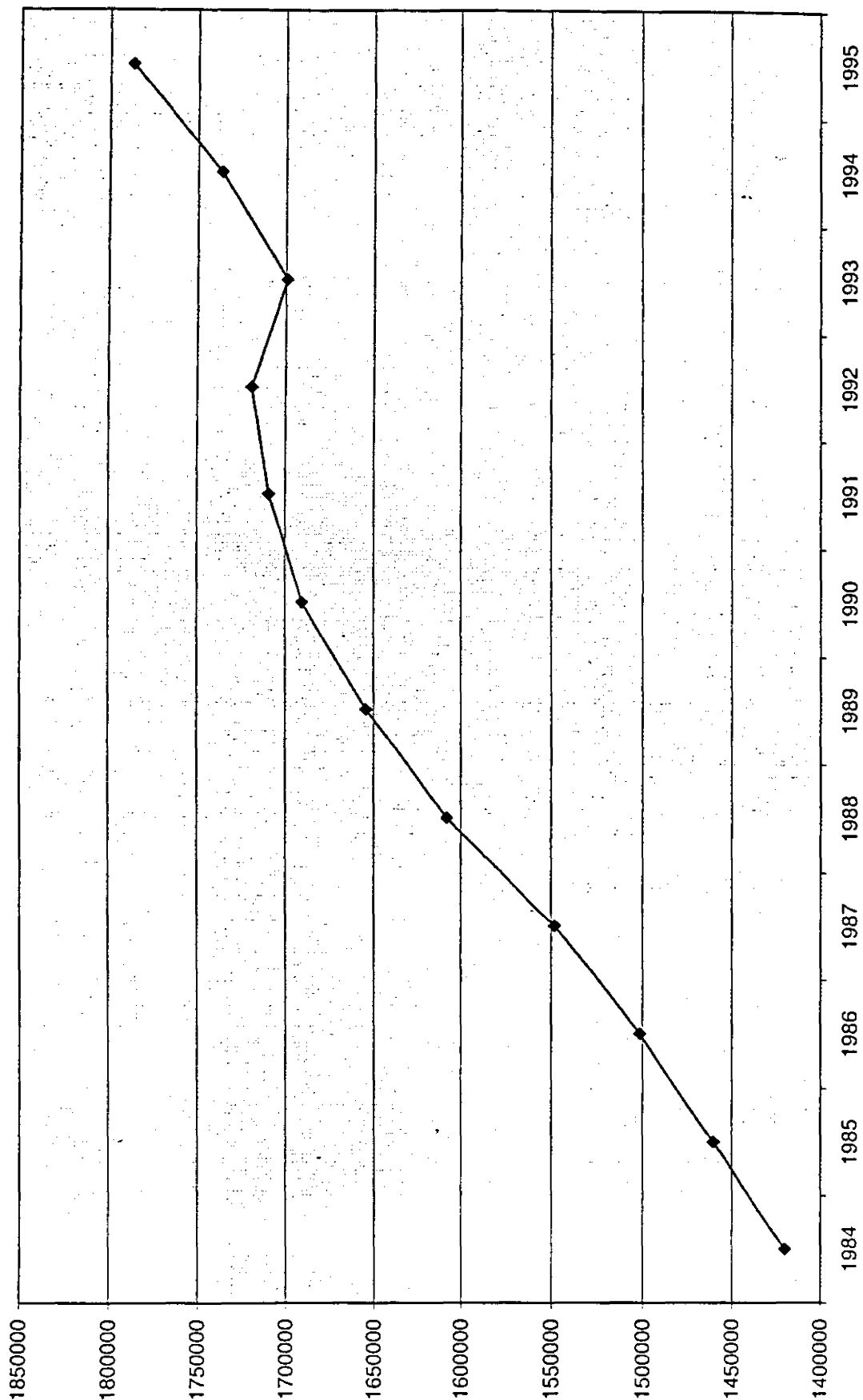


Figure 1.4: Youth Unempl. in Italy and Germany (age 15-24)

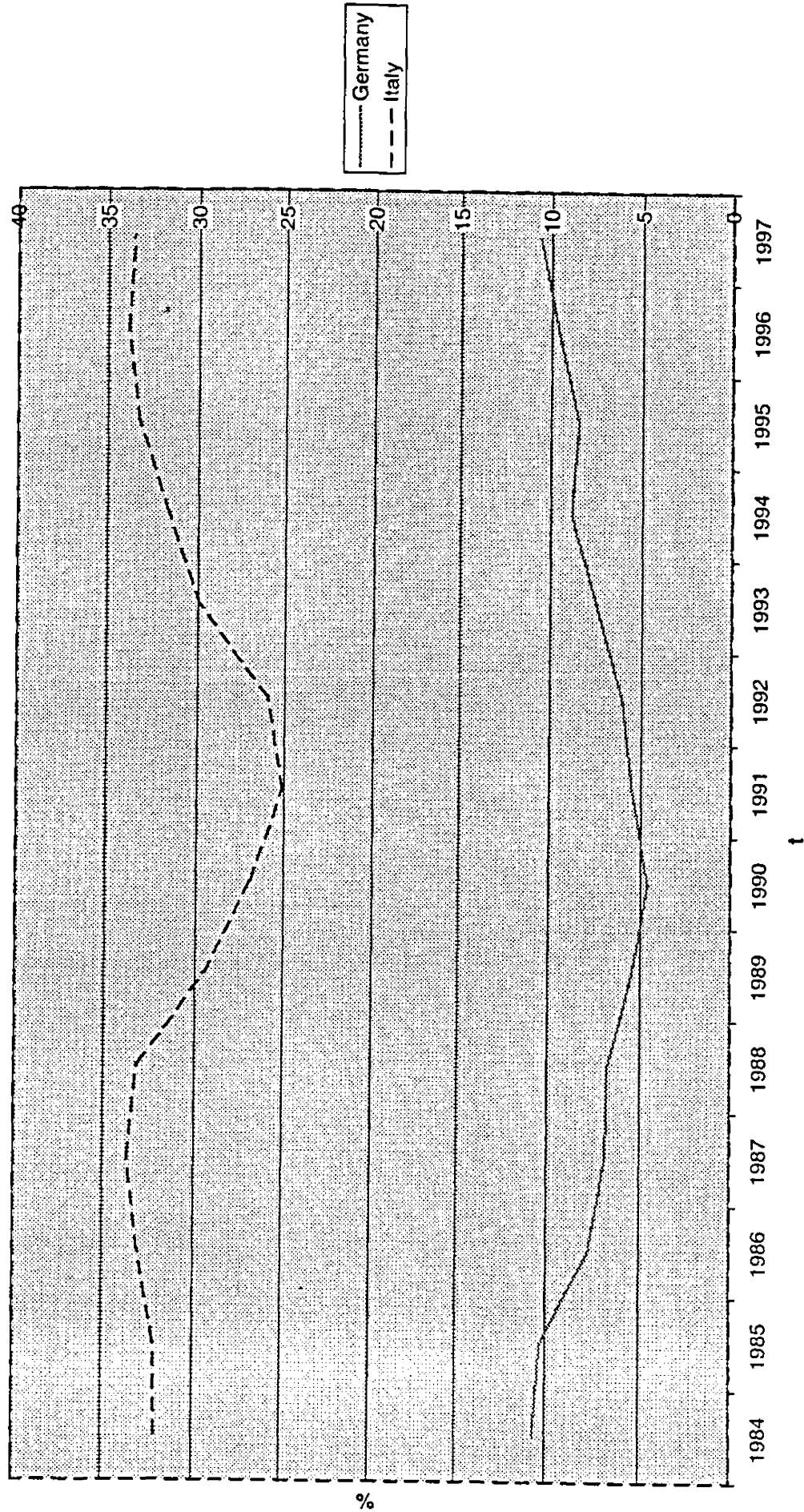
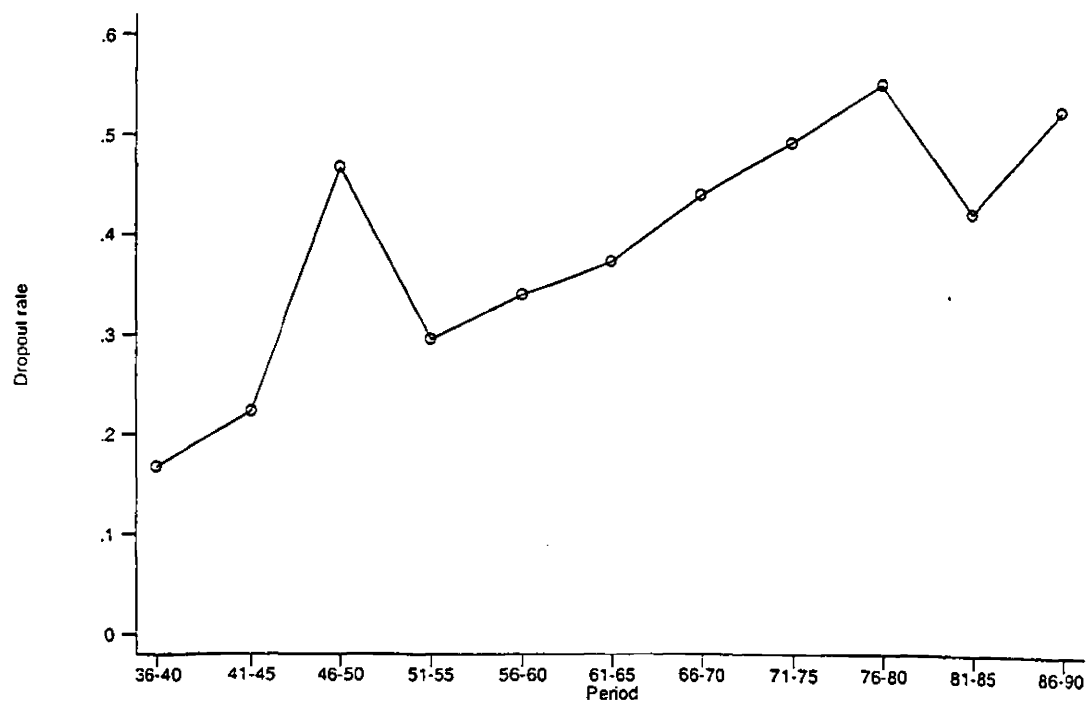


Figure 1.5: Historic Dropout Rates

Note: the figure displays the fraction of students enrolling in university in the given period that drop out without obtaining a degree.

Source: Indagine longitudinale sulle famiglie italiane (ILFI), own calculations.

Figure 1.6: Possible cases in the dropout model

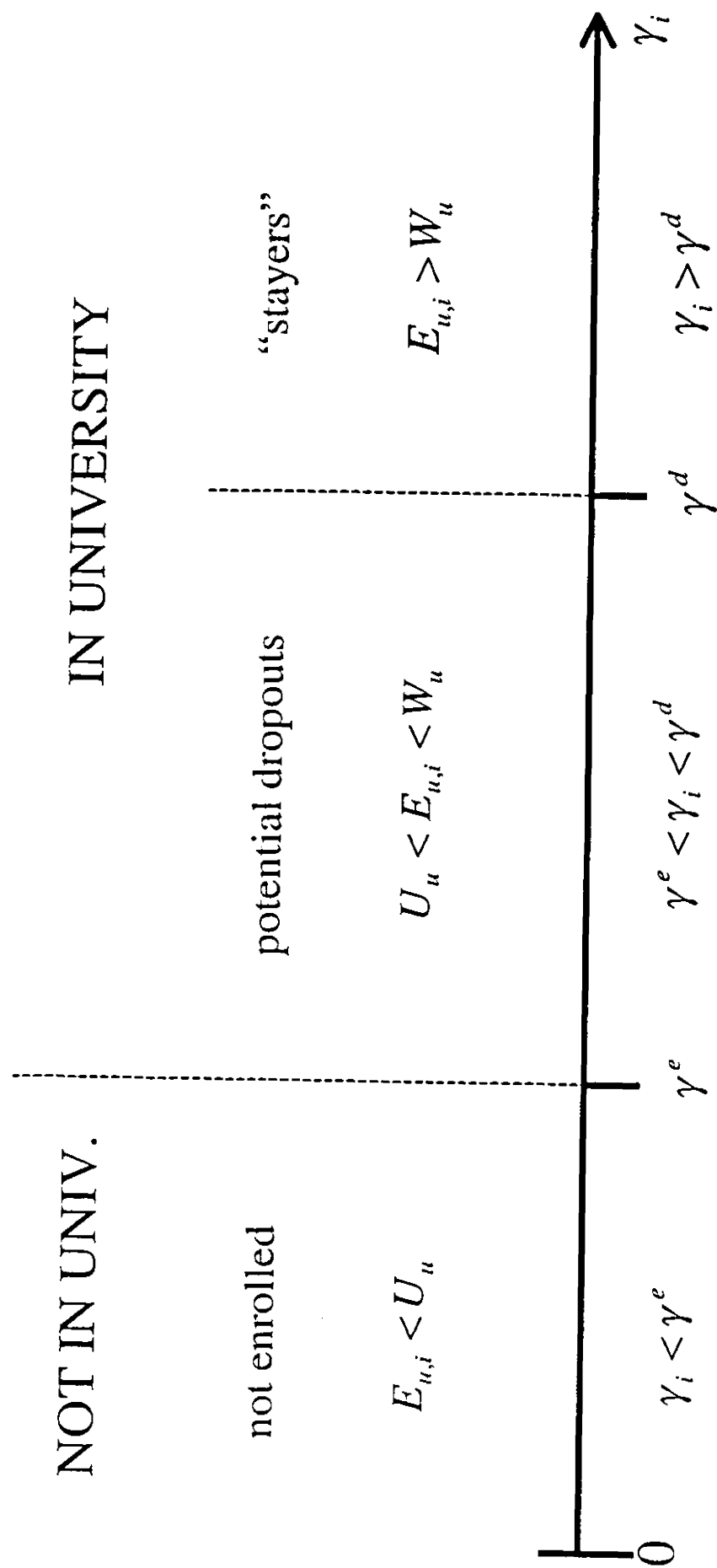
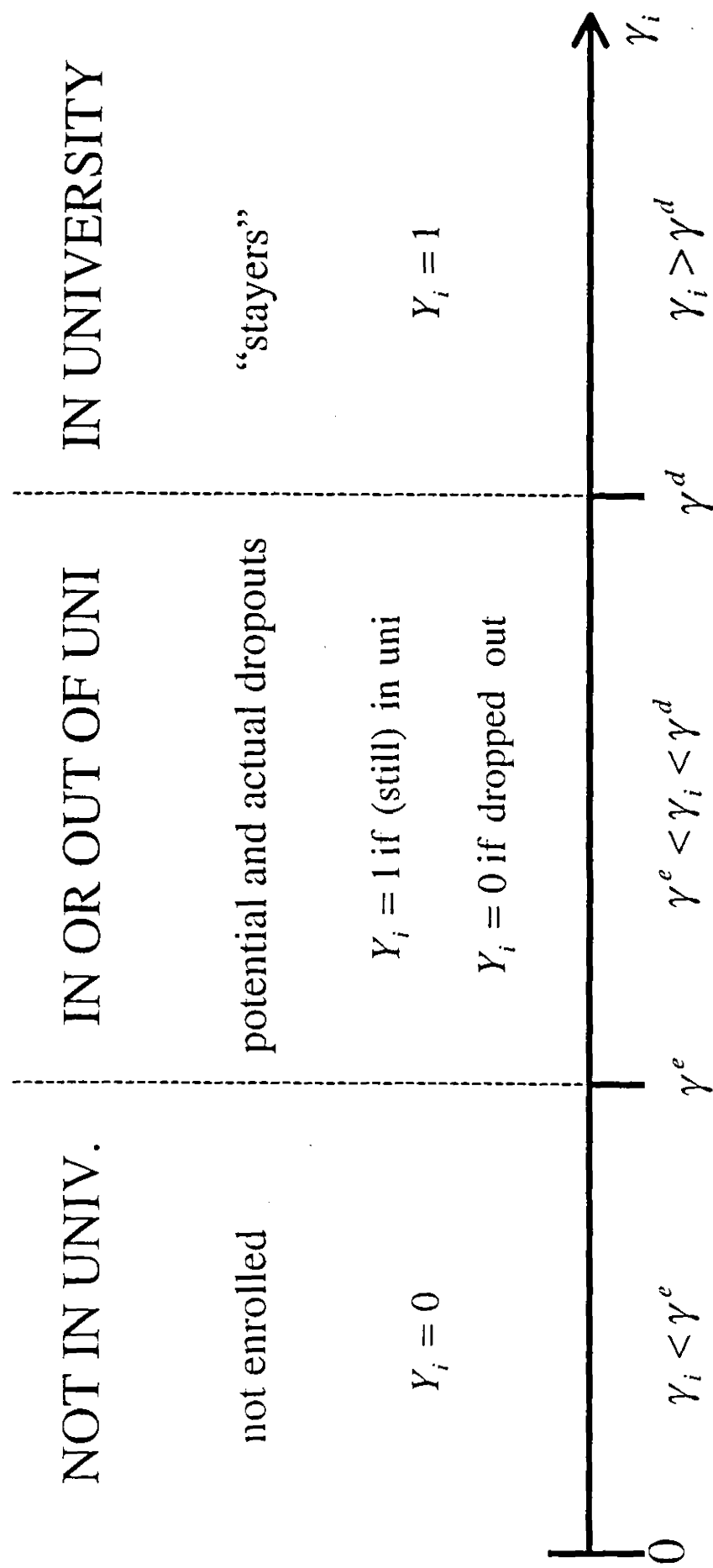


Figure 1.7: From the Model to the Data



Bibliography

- [1] Adam, Paula, and Patrizia Canziani. "Partial De-Regulation: Fixed-Term Contracts in Italy and Spain", *Centre for Economic Performance Discussion Paper no. 386*, March 1998.
- [2] Becker, Gary S. *Human Capital - A Theoretical and Empirical Analysis with Special Reference to Education*, 3rd ed. University of Chicago Press (for National Bureau of Economic Research), 1993.
- [3] Becker, Sascha O., Samuel Bentolila, and Andrea Ichino. "Parental Job Insecurity and Children's Emancipation: The Italian Puzzle", *mimeo*, 2001.
- [4] Brandolini, Andrea. "The Distribution of Personal Income in Post-War Italy: Source Description, Data Quality, and the Time Pattern of Income Inequality", *Temi di discussione, no. 350, Banca d'Italia* (http://www.bancaditalia.it/pubblicazioni/temidi/tema_350_99.pdf), 1999.
- [5] Card, David. "Earnings, Schooling and Ability Revisited", in: Solomon Polachek, ed., *Research in Labor Economics*, Vol. 14 (1995): 23-48.
- [6] Card, David. "The Causal Effect of Education on Earnings", in: O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3 (1999): 1801-63.
- [7] Card, David, and Lemieux. "Adapting to Circumstances: The Evolution of Work, School, and Living Arrangements Among North American Youth", *NBER WP 6142*, 1997.
- [8] Card, David, and Thomas Lemieux. "Dropout and Enrollment Trends in the Post-War Period: What Went Wrong in the 1970s?", *NBER WP 7658*, 2000.
- [9] Chiswick, Barry R. *Income Inequality: Regional Analyses within a Human Capital Framework*. New York: Columbia Univ. Press (for NBER), 1974.

- [10] *DTC Desktop Companion to the German Socio-Economic Panel Study (GSOEP) Version 2.2* - September 1998, Haisken-De New, John P. and Joachim R. Frick (eds.)
- [11] Flabbi, Luca "Returns to Education in Italy: OLS, IV and Gender Differences" *mimeo*, 2000.
- [12] Flinn, Cristopher. "Cohort size and schooling choice", *Journal of Population Economics*, Vol. 6, 1993: 31-55.
- [13] Griliches, Zvi. "Economic Data Issues", in: Z. Griliches and M.D. Intriligator, eds., *Handbook of Econometrics*, Vol. 3 (1986): 1465-1514.
- [14] Manacorda, Marco "The Evolution of Returns to Education in Italy, 1978-1992: Changing Institutions or Changing Supply and Demand ?" in Richard B. Freeman (ed.) *Inequality Around the World*, Macmillan & St. Martin's Press, 2000.
- [15] Manacorda, Marco, and Enrico Moretti. "The Evolution of Living Arrangements Among Young Italian Men" *mimeo*, 2001.
- [16] Pistaferri, Luigi. "Informal Networks in the Italian Labor Market", *Giornale degli Economisti e Annali di Economia*, vol.58 no. 3-4 (December 1999): pp. 355-375.
- [17] Pugliese, Enrico. "Special Measures to Improve Youth Unemployment in Italy" in McFate, Katherine, Roger Lawson, and William Julius Wilson, eds. *Poverty, inequality, and the future of social policy: Western states in the new world order*. New York: Russell Sage Foundation, 1995, pp. 439-60
- [18] Soro-Bonmati, Asuncion. "Labour Market Transitions and Transitions of Youth in Germany and Italy", *International Journal of Manpower*, vol. 21 no. 3/4 (2000): 206-226..
- [19] Wachter, Michael L., and William L. Wascher. "Leveling the Peaks and Troughs in the Demographic Cycle: An Application to School Enrollment Rates", *Review of Economics and Statistics*, Volume 66, Issue 2 (May, 1984): 208-215.
- [20] Wagner, Gert G., Richard V. Burkhauser, and Friederike Behringer. "The English Language Public Use File of the German Socio-Economic Panel." *Journal of Human Resources* (1993): 429-33.

Chapter 2

Returns to Education in Germany

(written with Frank Siebern-Thomas)

2.1 Introduction

The presence of heterogeneity in returns to schooling seems by now well established. Building on Gary Becker's (1967) model of optimal schooling according to which individuals choose their optimal schooling level by equating marginal benefits from continuing in education with the related marginal costs, recent theoretical contributions by *inter alia* Card (1995a,1995b) and Lang (1993) argue that individuals with different unobservable characteristics like ability, liquidity constraints or discount rates are likely to incur different marginal costs and benefits of further education and hence self-select into specific schooling levels. Such differences in the marginal costs and benefits of schooling imply different returns to schooling at different optimal schooling levels.

This in turn suggests the estimation of the returns to schooling on the basis of adequate instrumental variables and an interpretation of these estimates as *local average treatment effects* (LATE) along the lines of Angrist, Imbens and Rubin (1996): The estimated returns apply only to those individuals who are affected by the underlying instrument, i.e. those who only continue in school one more year because of their being induced to it by the instrument; in the language of the evaluation literature, the instrument is interpreted as an assignment to treatment, with one more year of schooling being interpreted as treatment. Moreover, different instruments will naturally affect different subgroups and hence lead to varying estimates of *the* returns to schooling.¹

¹For authoritative overviews of the recent literature on the identification and estimation of causal effects

Several empirical studies on the returns to schooling in the US seem to corroborate the LATE interpretation of instrumental variables estimates.² Card (1995b) and Kling (2000) e.g. use an indicator of the presence of a college in the county of residence at schooling age as instrument. They argue that this "college proximity" might allow individuals from low income (probably even liquidity constrained) families to attend college who otherwise (i.e. if they would have had to move to another county in order to go to college) would not have done so.³

For Germany, Ichino and Winter-Ebmer (1999, 2000)⁴ are the only authors we know of that provide LATE estimates of the returns to schooling.⁵ IWE (1999) especially contrast estimates obtained on the basis of two different instruments: first, parental educational background, and second, an indicator of the father's serving in the military during World War II. Since parental education as assignment mechanism is likely to affect less able children from well-off families (IWE call them the "stupid rich") the corresponding IV estimate is interpreted as a lower bound of the returns to schooling in Germany. On the other hand, "childhood during war" and particularly "father in war" are considered an extreme form of liquidity constraints that might hinder highly talented children from poor families (the "smart poor") to continue schooling. For this reason, the authors interpret the IV estimate based on the "war instrument" as an upper bound of the returns to schooling in Germany.

In this paper, we extend the IWE (1999)-study in several ways: First, we replicate the Card- and Kling-studies for Germany, making use of an instrument similar to Card's "college proximity". Second, we allow for a variable treatment intensity and try to characterize both the affected subgroups as well as the response functions, and third, we compare results for 1985 with those for 1995, thus testing indirectly for changes in the returns to schooling, in the instrument effectiveness, and in the response functions over time.

The results obtained on the basis of GSOEP data suggest that similar to the US results by Card and Kling, IV estimates of the returns to schooling are substantially higher

in economics cf. Angrist and Krueger (1999) and Card (1999).

²Further empirical studies on the returns to schooling in a LATE framework are e.g. Angrist (1990), Angrist and Krueger (1991), Angrist and Imbens (1995), and Kane and Rouse (1993).

³Cf. overview of IVE results by Card (1999).

⁴IWE (1999) draws on IWE (2000) where in addition to the instrument 'father in war' an indicator of the individual's having been in the age group 9-15 during the Second World War is used as an instrument. The latter paper is more specifically concerned with the long-run educational cost of World War II, while the first paper is more methodological and aims at providing evidence for heterogeneity in the returns to schooling.

⁵Lauer and Steiner (2000) do actually seem to follow a similar approach but they refrain from interpreting their estimates as local average treatment effects.

than corresponding OLS estimates. We show that individuals from disadvantaged family backgrounds profit most from a better schooling infrastructure prevalent in urban areas.

The remainder of the paper is organized as follows: in the next section we present Becker's model of optimal schooling. In section 3, we present some basic evidence on the relationship between educational attainment and college proximity using regional and GSOEP data. In section 4, we present the GSOEP data and describe our sample. Section 5 discusses the use of IV estimation in our context and presents the instruments used in the empirical analysis. There, we also summarize the results of this analysis and discuss their interpretation. Section 6 concludes.

2.2 Theoretical considerations

In this section we shortly recall Becker's (1964) model of endogenous schooling in the version laid out by Card (1995b). It provides both the rationale for heterogeneous returns to schooling and the basis for the LATE interpretation of our results.

An individual maximizes

$$U(y, S) = \log y - \phi(S) \quad (2.1)$$

where y is average earnings per year, S is years of schooling and $\phi(\bullet)$ is the cost of schooling. An individual's opportunities are represented by $y = g(S)$.⁶ The first order condition of the optimization problem is

$$\frac{g'(S)}{g(S)} = \phi'(S) \quad (2.2)$$

Now, assume for simplicity that

$$\frac{g'(S)}{g(S)} = \beta_i(S) = b_i - k_1 S \quad (k_1 \geq 0) \quad (2.3)$$

and

$$\phi'(S) = \delta_i(S) = r_i + k_2 S \quad (k_2 \geq 0) \quad (2.4)$$

⁶There is considerable discussion in the literature as to which variable best describes the theoretical concept of human capital. Griliches (1977) points out that years of schooling is rather one of the inputs of the human capital production process than its outcome. To the extent that output measures are unavailable, years of schooling as a proxy for human capital is the best variable we can get to describe what is valued in the labor market.

The optimal schooling level is then given by $S_i^* = (b_i - r_i)/k$, where $k = k_1 + k_2$. Integrating out (2.3) yields

$$\log y = b_i S - 0.5 k_1 S^2 \quad (2.5)$$

Equations (2.3) and (2.4) clearly state the reason for heterogeneous returns to schooling: Individuals are likely to differ in either marginal costs r_i or marginal benefits b_i and are therefore likely to choose different optimal schooling levels as shown in figure 2.1.

This is exactly what is exploited by the LATE-IV approach. A given instrument will affect different margins, i.e. different subpopulations at different schooling levels. As explained in detail in Angrist and Imbens (1995) we can hope to estimate only the average marginal return to schooling for a well-defined subgroup which is affected by the instrument. In the presence of heterogeneity, the notion of a unique return to schooling is hence nonsensical. In section 2.5.2 we are going to explain this in further detail.

We actually estimate the following system of equations:

$$y = X\beta + S\gamma + \varepsilon \quad (2.6)$$

$$S = X\delta + Z\alpha + \eta \quad (2.7)$$

where Z is an instrument or set of instruments. For the LATE interpretation of IV to apply to the estimate of γ in (2.6), the conditions in Imbens and Angrist (1994) have to apply.⁷ This approach thus makes a good out of the two main problems faced in a simple OLS regression of (2.6): the problem of self-selection into schooling and heterogeneity in returns to schooling. The main problem in empirical applications is, of course, to find an adequate instrument as an exogenous source of variation in education choices.

2.3 Educational outcomes and returns to schooling in Germany

In this section, we present descriptive evidence based on regional data for some recent years (1996-1998).⁸ We collected data about school completion rates and school infrastructure as

⁷Further assumptions implicit in equations (2.6) and (2.7) are log-linearity of earnings in schooling and the absence of degree effects (*sheepskin effects*). See Card (1999) for empirical evidence on the absence of sheepskin effects in the US.

⁸The data had to be obtained from the single regional statistical offices (*Statistische Landesämter*) because to our knowledge no consistent educational data base exists at the national level.

well as some information about the state of the labor market at the level of counties (*Kreise*). These data show, in particular, a huge variation in completion rates across counties as well as a positive correlation between completion rates and schooling infrastructure.

2.3.1 Some background information using regional data

High school completion rates (*Abitur*) in Germany range from roughly 8% (in the *Südwestfalz*) to 52% (in *Darmstadt*) of all school leavers across counties and hence show astonishingly strong regional variation. To see whether there is any systematic relationship between these high school completion rates on the one side and the schooling infrastructure on the other side, we plotted the percentage of school leavers having *Abitur* against the log of the number of high schools per square kilometer as a measure of schooling infrastructure (see figure 2.2). The availability of high schools is in fact seen to be highly correlated with high school completion rates.⁹

A higher average distance to the nearest high school is likely to increase the costs of education. Apart from the (time) opportunity costs of having to travel more, direct costs involve additional transport costs. All other costs do a priori not differ by distance to school. They might differ, however, across the various German regions (*Länder*) which are solely responsible for educational matters. Although there are generally no school fees neither for primary and secondary schools nor for universities, regulations regarding the public provision of books and other material used by students or subsidies for book purchases to low income families as well as regarding transportation subsidies for students do actually differ significantly across the various *Länder*. In many regions subsidies to either transport or book purchase are limited to students up to compulsory school age (i.e. 18 years old) or some other specific age (15 or 16 years old) and have to be borne fully by older students. Last but not least, the schooling years necessary for high school completion amount to 13 years in the West German *Länder* and Brandenburg as opposed to only 12 years in the remaining new German *Länder*. At university, the only fee to pay is for social security and health contributions.¹⁰

To sum up, using regional data we find lower high school completion rates in rural, less densely populated regions with a poorer schooling infrastructure. In addition, using microdata (GSOEP) we find lower high school completion rates for individuals who grew up

⁹Of course, this is not necessarily a causal relationship driven by the supply of high schools. It could also be that lower demand for higher education causes less supply by the state.

¹⁰In the later regressions, we try to capture differences in regulations across states by including a set of state dummies.

in rural as opposed to urban areas (see table 2.1).¹¹

Average years of schooling by agglomeration show a similar pattern as can be seen from table 2.2.

Do these differences tell us something about regional variations in the quality of schools and/or high school degrees (as often suggested in the political debate) or are they indicative of regionally varying opportunity costs related to longer schooling?

Our conjecture is that higher costs of education in regions with 'poor schooling infrastructure' reduce private investments in schooling, at least among children from relatively low-income/high discount rate families. This is also suggested by existing empirical studies on the returns to schooling based on instrumental variable estimation (Card (1995b), Kling (2000)). Card finds that the IV estimates of the earnings gain per year of additional schooling (10-14%) are substantially above the earnings gains estimated by a conventional OLS procedure (7.3%). Kling (2000), using Card's data, confirms Card's results and further characterizes the group of students affected by differences in place of childhood.

2.3.2 Previous studies

Previous results for Germany are based on simple OLS regressions of earnings on schooling. Using years 1984 and 1985 of the German Socioeconomic Panel (GSOEP), Wagner and Lorenz (1989) estimate returns to schooling of 6.5%. In a further study Lorenz and Wagner (1993) give a range of 6.2-7.0% based on the Luxemburg Income Study (LIS 1981) and of 4.0-4.9% using data of the International Social Survey Program (ISSP 1987).

To our knowledge, the only studies using IV estimation are Ichino and Winter-Ebmer (1999, 2000) and Lauer and Steiner (2000). The former authors exploit three different instruments: an indicator of father's education, an indicator of whether an individual was 10 years old during World War II and an indicator of whether their father was in war in this period. Using data from the GSOEP (1986), they give a lower bound of 4.8% and an upper bound of 14% for the return to schooling for those subpopulations that are affected by the respective instruments. The latter authors not only estimate the returns to schooling using various estimation methods but also employ IV estimators on the basis of a whole long list of different instruments. They are above all interested in an analysis of the robustness of the estimated returns to schooling with respect to the various estimation methods and

¹¹Using regional data and defining agglomerations by quartiles of population density - which obviously do not coincide with the GSOEP classification - we observe a similar pattern. Going from the most densely to the least densely populated quartile, high school completion rates in 1997 are 30.92, 23.07, 18.86, and 19.40 respectively.

do not provide a LATE interpretation of the obtained IV estimation results. Moreover, the authors conclude that there is no statistical evidence for heterogeneous returns to schooling with respect to unobservable characteristics.

2.4 Data description and sample selection

We use the German Socioeconomic Panel because it is the richest German micro data set to study our question of interest. It contains information on education, income, personal and job characteristics, family background, and biography information that we exploit in the IV estimation below.

We only keep the full-time employed in 1985 or 1995 who have no missing information on our variables of interest, in particular labor income and schooling. In tables 2.3, 2.4, and 2.5, we show descriptive statistics for the dependent variables, for schooling variables, and for exogenous variables. As for schooling variables, we present both average years of schooling along with degree information.

2.5 IV Estimation of the Returns to Education

2.5.1 Choice of instrument

Previous studies have used a broad range of instruments to establish causality in the returns to schooling (see Card, 1999) and the references therein). The choice of an instrument has several important aspects. First, econometrically speaking the instrument should fulfill the exclusion restriction, i.e. have an effect on earnings only via the schooling channel but no direct effect on earnings. Second, heterogeneity in marginal costs and benefits of schooling and therefore the absence of a unique return to schooling for the population as a whole can be exploited by choosing an instrument which describes a quasi-experiment of important policy interest. So, IV estimation is not just the solution to the econometric problem of possibly biased OLS estimates but allows to analyze interesting policy questions. On the basis of these two considerations, we choose our instrument 'place of childhood' which is similar to Card's (1995b) college proximity indicator. It has not yet been used for German data and allows us to address the question as to *who* profits *how* from differences in schooling infrastructure across different places of childhood.

The question on place of childhood in the GSOEP questionnaire is expressed as follows:

"Did you spend the major portion of your childhood up to age 15 in a) a city, b) a big town, c) a small town, or d) in the countryside?"

In the sequel, we are going to use three different binary indicators based on this question: 'spent childhood in a city' (pc1), 'spent childhood in a city or big town' (pc2), and 'spent childhood in an urban area' (pc3), i.e. in a city, or in a small or big town. Table 2.6 shows the percentage of the sample with given instrument status.

2.5.2 Which Effect Can We Identify? The variable treatment approach to the estimation of returns to schooling

The IV estimate of the returns to schooling based on 'place of childhood' as an instrument identifies a causal effect for well-defined subpopulations and schooling levels. The implied natural experiment uses place of childhood as assignment to treatment (Z), the schooling level as treatment (S), and $\log(\text{monthly earnings})$ as outcome (Y).

The model we estimate is an extension of Rubin's Causal Model (RCM) to variable treatment intensity. Assume that each individual would earn Y_j if he or she had j years of schooling for $j = 0, 1, 2, \dots, J$. The objective is to uncover information about the distribution of $Y_j - Y_{j-1}$, which is the causal effect of the j th year of schooling. This will help us understand under which conditions and for which subpopulation of interest γ can be given a causal interpretation. In general, estimates of γ in equation (2.6) have a causal interpretation only if they have probability limit equal to a weighted average of $E[Y_j - Y_{j-1}]$ for all j in the subpopulation of interest.

We can define potential schooling levels and potential outcomes for all potential values of the instrument (e.g. grown up in the countryside, in a small town, in a town, in a big city) for each individual. We define $S_Z \in \{0, 1, 2, \dots, J\}$ to be the number of years of schooling completed by a student conditional on the values of the instrument. Let's initially assume that Z is coded to take on only two values, 1 and 0, indicating that the place of childhood was either in an urban area or in the countryside. S_1 then denotes the years of schooling that would be obtained by an individual growing up in an urban area, and S_0 is the years of schooling of the same individual if he or she grew up in the countryside. In the data, for each individual we observe the triple (Z, S, Y) , where Z denotes the place of childhood, $S = S_Z = Z * S_1 + (1 - Z) * S_0$ is years of completed schooling, and $Y = Y_S$ is earnings.¹²

¹²Note that, for simplification, we do not use distinct notation for random variables and observations. More correctly, we should denote observations as $(Z_{obs}, S_{obs}, Y_{obs})$, where Z_{obs} denotes the observed place of childhood, $S_{obs} = S_{Z_{obs}} = Z_{obs} * S_1 + (1 - Z_{obs}) * S_0$ is observed years of completed schooling, and $Y_{obs} = Y_{S_{obs}}$.

The main identifying assumption is the following

Assumption 1 (*Independence*)

The random variables $S_0, S_1, Y_0, Y_1, \dots, Y_J$ are jointly independent of Z .

In our case this requires that place of childhood has no effect on earnings other than through its effect on schooling. This implies the existence of unit-level causal effects. To identify a meaningful average treatment effect, the literature typically assumes a constant unit treatment effect, $Y_{ij} - Y_{i,j-1} = \alpha$, for all schooling levels j and all individuals i . Angrist and Imbens (1995), however, impose a nonparametric restriction on the process determining S as a function of Z instead of restricting treatment effect heterogeneity. They impose the following

Assumption 2 (*Monotonicity*)

With probability 1, either $S_1 - S_0 \geq 0$ or $S_1 - S_0 \leq 0$ for each person.

Angrist and Imbens (1995) further show that for multivalued treatments ($J > 1$), assumption 2 has the testable implication that the cumulative distribution function (CDF) of S given $Z = 1$ and the CDF of S given $Z = 0$ should not cross.

From the above assumptions follows the main result in the framework of multivalued treatments:

Theorem 1 *Suppose that Assumptions 1 and 2 hold and that $\Pr(S_1 \geq j > S_0) > 0$ for at least one j . Then*

$$\frac{E[Y|Z=1] - E[Y|Z=0]}{E[S|Z=1] - E[S|Z=0]} = \sum_{j=1}^J \omega_j \cdot r(j) \equiv \gamma \quad (2.8)$$

where

$$\omega_j = \frac{\Pr(S_1 \geq j > S_0)}{\sum_{i=1}^J \Pr(S_1 \geq i > S_0)} \quad (2.9)$$

denotes weights and where the response function is defined as

$$r(j) \equiv E[Y_j - Y_{j-1} | S_1 \geq j > S_0] \quad (2.10)$$

is observed earnings as a function of observed schooling.

This implies that $0 \leq \omega_j \leq 1$ and $\sum_{j=1}^J \omega_j = 1$, so that γ is a weighted average of per-unit average causal effects along the length of an appropriately defined causal response function. Angrist and Imbens (1995) refer to the parameter γ as the *average causal response* (ACR).

The *ACR weights* ω_j are proportional to the number of people who, because of the instrument, change their treatment from less than j units to j or more units. The *response function* $r(j)$ gives the average difference in the outcome for those who change their treatment from less than j units to j or more units. In the case of further covariates, the analysis is slightly more complicated and requires weighting by the conditional variance of Z .

In our example, IV generates an estimate of the average causal effect among individuals with different marginal benefits from schooling: First, different subgroups are affected by different instruments. Second, individuals in these subgroups are affected by the respective instrument in different ways. And third, the instrument may induce changes of behavior at different levels of schooling.

In the empirical part, we present both the weighting function and the response function for the given choice of instrument and thereby try to characterize the affected subgroups and schooling levels.

2.5.3 IV Estimation Results

We started by estimating an OLS regression of earnings on years of schooling controlling for sex, experience and tenure on the job polynomials, yielding estimates in the usual range of 6.7% and 6.6% for 1985 and 1995 respectively.

For the reasons given above, these estimates are probably not amenable to an interpretation as the causal effect of schooling on earnings. We therefore performed an IV estimation of the returns to education on the basis of the instruments suggested above. The instrumental variables estimates of the returns to schooling on the basis of the chosen instrument have been computed using the two-stage least squares procedure: in the first stage, the years of schooling are regressed on the whole list of exogenous variables augmented by the respective instrumental variable using a simple linear probability model; in the second stage, the predicted value of the dependent variable from the first stage regression is then used as additional regressor in the outcome equation instead of the schooling years itself. Table 2.7 contains the IV estimation results for the various chosen instrumental variables. Further, first-stage t -statistics and partial R^2 measures are reported as a diagnostic tool following the suggestions of Bound et al. (1995) and Staiger and Stock (1997). In all cases, the instrument quality seems reasonable as suggested by these measures.

The returns estimated using either of these instruments are considerably higher than the

OLS estimates. In 1985, the point estimates are 12.6%, 12.5% and 13.3% for the binary instruments 'spent childhood in a city', 'spent childhood in a city or big town', and 'spent childhood in an urban area'. A similar picture arises in the 1995 data. Throughout, the IV estimates are nearly double the size of the OLS estimates. In the light of the LATE framework, these results can be interpreted as the returns to education for those who acquired more education *because* they are living in an area with a good schooling infrastructure.

2.5.4 Internal validity of the instruments

To check the internal validity of the instrument for identification of the LATE parameter, we have to check the assumptions given in Angrist, Imbens and Rubin (1996).¹³ Not all of these assumptions are in general rigorously testable. We can only argue and give corroborating evidence as we do in the sequel.

We can be quite confident that the *SUTVA assumption* is satisfied in our sample. It requires potential earnings to be unrelated to the amount of schooling taken by *other* individuals in the sample. This assumption is more likely to be violated in clustered samples.

Strongly ignorable assignment to treatment requires that after controlling for observable characteristics, unobservables like ability should be randomly distributed across different places of childhood. This assumption could be violated if parents endogenously choose to live in an urban area because of better schooling infrastructure. Most of this potential selection into places of living is probably controlled for by observables. In any case, geographical mobility in Germany is quite low by international standards. While Germany has 16 states and about 80 million inhabitants, the US have 51 states and about 250 million inhabitants, so average population per state is relatively similar, the US states being bigger in size, however. While in the US, 3% of the population move across state borders every year, in Germany only 1% of the population move across state borders.¹⁴ Not only are mobility rates low anyway,

¹³ AIR (1996) prove that the instrumental variables estimate of δ in the heterogeneous *treatment* effect model has a *causal* interpretation as *local average treatment effect* under the following assumptions:

(1) Potential *outcomes* for each individual i are unrelated to the *treatment* status of other individuals. (*stable unit treatment value assumption (SUTVA)*)

(2) Conditional on observables, the *assignment to treatment* is random. (*strongly ignorable assignment to treatment*)

(3) The *treatment* probability is a nontrivial and monotonous function of the instrument, i.e. $E[D_{i1} - D_{i0}] > 0$. (*strong monotonicity*)

(4) The (unit-level) potential *outcome* variables depend on the assignment status Z_i only through the *treatment* status D_i , i.e. $(Y_{i0}, Y_{i1}) \perp Z_i | D_i$. (*exclusion restrictions*)

¹⁴ Data come from the US Census Bureau website

but the reasons for moving are very unlikely to be related to schooling infrastructure as well. The GSOEP data contain a question on reasons for move. In 1997, respondents can give a maximum of three out of a list of 15 possible reasons. Overall, 8.6% of the movers give "other family reasons" (i.e. not divorce, marriage and leaving parent's home) as reason for move. If at all, families that move to give their kids access to a better schooling infrastructure might show up in this group. For families with kids under age 18 (i.e. those families for whom schooling infrastructure might play a role), the percentage moving for "other family reasons" is even lower yet (5.2%), thus making "better schooling infrastructure" an even more unlikely reason for moving. We conclude that our estimates are very unlikely to suffer from violation of the *strongly ignorable assignment to treatment* assumption.

Strong monotonicity compares again two counterfactual situations: an individual growing up in a city (i.e. in region with good schooling infrastructure) takes at least as much schooling as if he had grown up in the countryside (i.e. in a region with a worse infrastructure). This assumption rules out *defiers*, i.e. individuals who, if growing up in a city, take less schooling than if growing up in the countryside. In theory, there might be individuals who take less schooling growing up in an urban area due to e.g. drugs and delinquency, but growing up in a rural area would have obtained more schooling. In a similar way, labor demand in cities might be higher and therefore students might have more outside options in a city as compared to an urban area and for some individuals these outside options might lead to a lower schooling level. While we cannot really rule out that there are some cases like this, for the reliability and interpretability of our estimates it is important that the fraction of defiers is nevertheless very small. One testable implication of strong monotonicity is that the cumulative density functions of schooling by instrument status do not cross. As we will show, this holds in our data and makes us confident that violation of the strong monotonicity assumption is not a serious issue here.

The *exclusion restriction* would be violated if there existed a direct effect of the suggested instrument on earnings, e.g. in the form of an 'urban wage premium'. We are in the fortunate situation to have some information about the current place of living. The GSOEP data contain both current state (*Bundesland*) of residence as well as the so-called Boustedt regions.¹⁵ We find that by including these further controls, in 1985 the estimated returns to schooling do not change and in 1995 they even go slightly up. When controlling for state

(<http://www.census.gov/population/socdemo/migration/tab-a-1.txt>) and from the website of the German National Statistical Office

(<http://www.statistik-bund.de/jahrbuch/jahrstab5.htm>).

¹⁵Boustedt (1970) classifies urban regions into seven categories, assigns the neighbouring communities of an urban center to four different sub-categories from "rural" to "urban center".

dummies, the coefficients on the Boustedt dummies are found to be statistically insignificant. We might therefore conclude that there is no violation of the exclusion restriction through an urban wage premium.

Another reason why the exclusion restriction might be violated is that school quality might vary by place of childhood. In this case, controlling for characteristics of the current place of living is not sufficient because people might have moved and the decision to take further schooling depended on their place of childhood and not on their current place of living. To see if this is a valid objection, we follow an idea similar to Card (1995b) and Kling (2000). They propose to define family background quartiles across which the returns to schooling will vary. In order to test whether college proximity is a legitimate instrument, they use the interaction of college proximity with an indicator for low parental background as an instrument and control for the main effect of college proximity. Translated to our setup, the idea is that our instrument is unlikely to affect individuals from higher family background quartiles because they have the necessary support by their family to pursue further education even if the respective schools are not nearby. So, using the instrument as such or using the instrument interacted with an indicator of low family background is the same, and gives us one more degree of freedom, namely allows us to control for the main effect of the 'place of childhood' indicator. We will further discuss the construction of the family background quartiles in the following section. There, we also use them to characterise the subgroup of compliers, so they serve a double purpose.

Let us shortly summarize the results of the estimation using the interacted instruments. We find that indeed the main effect of 'growing up in an urban area' is small in size and statistically insignificant.¹⁶ The lower panel of table 2.7 shows that the point estimates are lower than the ones where we do not control for the main effect of 'growing up in an urban area', but that they are still considerably higher than the OLS estimates. On the basis of this evidence in favor of both the absence of urban wage premia and the validity of the exclusion restriction, we conclude that the returns to education for the subgroups of compliers, i.e. those individuals who only acquire more schooling when enjoying a good schooling infrastructure, are significantly and substantially higher than the simple OLS estimates. In the following section, we turn to the characterization of the subgroups affected by our instrument.

¹⁶The coefficients on the main effect $pc1$ is 0.012 with a standard error of 0.019 in 1985, and 0.011 with a s.e. of 0.020 in 1995.

2.5.5 External validity of the instruments

If we want to generalize our estimates to some larger populations ("external to the sample"), we have to characterize as closely as possible the subgroups affected by our instrument and the size of the effect on them. We suggested above that the effect of schooling infrastructure is more important for children from less advantaged family backgrounds. We follow Card and Kling in defining family background quartiles in the following way: First, we perform a regression of years of schooling on the subgroup of people who spent their childhood in a rural area. Then, based on the parameter estimates obtained, we predict - for all individuals - their 'counterfactual schooling level if they had grown up in a rural area' and split the sample into four quartiles, from the lowest (fbq1) to highest (fbq4).

Table 2.8 presents some summary statistics on average years of schooling by instrument status and family background quartile for the years 1985 and 1995. Apart from the fact that average years of schooling are higher for those who grew up in urban areas, the table clearly shows that for those who have a higher predicted (counterfactual) schooling level, also actual schooling attainment is higher.

Table 2.9 further shows the distribution of family background and individual variables across these 'counterfactual schooling quartiles'.¹⁷ There is no single individual in the lowest three family background quartiles whose father has a university degree. Conversely, there is virtually no individual in the two highest background quartiles who has a father without a schooling degree. We also see that a higher percentage of those in the upper family background quartiles did actually grow up in a city.

The IV estimate of the returns to schooling can be interpreted as a weighted average of the potentially differing treatment effects across the four background quartiles, γ_q , with the weight given to each quartile q by the product of the proportion of the population in that subgroup (w_q) and the impact on schooling for that subgroup (ΔS_q). This allows us to write

$$\gamma = \sum_{q=1}^4 \frac{w_q \Delta S_q \gamma_q}{\Delta S}$$

We give the weights w_q in table 2.10.

Table 2.11 shows the differences in schooling levels by instrument status for the population as a whole (ΔS).

Figures 2.3 and 2.4 further split up the information of table 2.10 by family background quartiles for 1985 and 1995 respectively. In 1985, the actual average education difference by

¹⁷It is interesting to note that in the lowest background quartile, none of individuals report that either their father or mother graduated from high school.

instrumental status is much larger for the two lower background quartiles, supporting the suggestion of section 5.3 that instead of our indicator for 'growing up in an urban area' we can equally well use this indicator interacted with poor family background. This allows us to use the main effect of 'growing up in an urban area' in the estimation and thereby control for there being an urban wage premium. We already reported the results of this exercise in the previous subsection.

2.5.6 Characterizing the response function

The response function can be estimated from the cumulative distribution functions (CDF) of schooling at different values of the instrument. The difference in the CDFs is equivalent to the fraction of the population who received at least one more year of schooling due to the instrument. Figure 2.5 shows the difference in the CDFs for the 1985 sample using *pc1* as an instrument.¹⁸ It indicates that schooling infrastructure has its largest effect at 11 years of schooling. More specifically we interpret the estimates to indicate that around 10 percent of individuals with similar demographics are induced to obtain more years of schooling due to better schooling infrastructure.

It is even more interesting to break down the response function by background quartiles. Figure 2.6 shows that the response function of the two lower background quartiles peaks at 10 years of schooling while the response of the two upper quartiles is concentrated among those with 13 or more years of schooling. Furthermore, the fraction of 'compliers' in the two upper quartiles is overall much lower, again showing that the instrument affects mainly the two lower family background quartiles.

From a policy point of view, this result suggests that the provision of schools beyond 10th grade, i.e. basically the provision of (senior) high schools (*Gymnasien*), can considerably increase the fraction of youths from disadvantaged backgrounds who obtain more schooling.

For 1995, the picture is slightly different. First, figure 2.7 suggests that for this later cohort, schooling infrastructure increased educational attainment at a later stage in educational careers.

Overall, in 1995 the response function is flatter and takes on lower values than in 1985. Second, breaking down by background quartiles, we find that the point of maximum response has moved to the right for all subgroups. Also has the fraction of the population in all subgroups who respond to our instrument decreased (see figure 2.8)

The fact that figures 2.5 and 2.7 display only non-negative values is equivalent to say-

¹⁸Figures based on the instruments *pc2* and *pc3* show a similar pattern and are therefore not shown here.

ing that the CDFs for $Z = 1$ and $Z = 0$ don't cross, a finding that supports the strong monotonicity assumption laid out in section 2.5.4.

To sum up, there seems to be a decreasing effect of our instrument on lower schooling levels and/or an increasing effect of the instrument on higher schooling levels. This also explains why returns to education seem to have decreased between 1985 and 1995.

2.6 Summary and conclusions

This study corroborates the general finding of other studies based on IV estimation that OLS estimates are downward biased. It confirms the empirical evidence that different instruments lead to different estimates of the schooling coefficient, underlining the fact that returns to schooling are heterogenous. Our estimates remain within the bounds given by IWE (1999). We find that individuals from 'poor family background' respond most strongly to the instrument 'place of childhood'. Their response is further most pronounced at low schooling levels whereas the response of individuals with 'rich family background' is most pronounced at higher schooling levels. Finally, this approach allows us to detect changes in the response function over time.

The temporal variation of returns to schooling operates through two different channels. First, temporal variation in the covariate weights leads to a reweighting of the returns for different subgroups. We conjecture that there is a decreasing fraction of compliers from a poor family background and/or an increasing fraction of compliers from a rich family background. Second, temporal variation of returns to schooling is also due to temporal variation in the response functions. There seems to be a decreasing effect of our instrument on lower schooling levels and/or an increasing effect of instrument on higher schooling levels.

The finding that educational attainment crucially depends on the provision of post-compulsory schooling in proximity to the place of living, has important policy implications. Consider the case of a regional government that has decided to devote a certain amount of money to the improvement of upper secondary schooling infrastructure.¹⁹ It then faces the decision *where* to build the school, in an urban area or in a rural area, or similarly whether to build one big school in a city or some smaller schools in the countryside. If the per student cost of providing further places at school is constant independent of where schools are built, our results clearly indicate that students living in areas with a less favourable

¹⁹We do not address the cost-benefit issue here, i.e. we do not ask whether for the region as a whole investing in schooling infrastructure is beneficial. In contrast, we take an individual-level perspective and take the provision of funds by the government as given in this thought experiment.

schooling infrastructure would probably benefit most from such an investment because of their above average marginal returns to education. To the extent that schooling infrastructure is correlated with the degree of urbanisation, providing a better schooling infrastructure especially in rural areas could thus considerably increase the incentives for individuals from disadvantaged family background to acquire more education and thus improve their long-run prospects in the labor market.

It is important to note, though, that the policy implication might be quite different for the case in which the federal government increases schooling infrastructure in the country as a whole. In this case there might be general equilibrium effects that decrease the return to education in the long run due to an overall higher supply of better-educated individuals (see Heckman et al., 1999). The policy implications of this paper do therefore refer to the optimal allocation of schools but not necessarily to the optimal overall spending on schooling infrastructure.

2.7 Appendix

Table 2.1: High school completion rates by agglomeration

	GSOEP	GSOEP
agglomeration	1985	1995
city	18.49	25.00
big town	15.76	20.12
small town	10.94	18.70
in the countryside	8.58	12.97

Table 2.2: Years of schooling by agglomeration

	GSOEP	GSOEP
agglomeration	1985	1995
city	12.01	12.42
big town	12.00	12.37
small town	11.28	11.97
in the countryside	11.10	11.63

Table 2.3: Summary statistics on outcome variables

Variable	N	Mean	Std.Dev	Min	Max
1985					
gross monthly income	4125	2983.33	1382.73	0	19000
net monthly income	4277	2029.74	959.23	0	13000
1995					
gross monthly income	3242	4524.49	3038.21	0	99999
net monthly income	3287	2984.71	1845.25	0	50000

Table 2.4: Summary statistics on education

Variable	N	Mean	Std.Dev	Min	Max
1985					
years of schooling	4617	11.47	2.78	7	19.5
Hauptschule	4617	0.43	0.50	0	1
Realschule	4617	0.34	0.48	0	1
Fachhochschulreife	4617	0.03	0.18	0	1
Abitur	4617	0.09	0.29	0	1
Apprenticeship	4617	0.64	0.48	0	1
University degree	4617	0.10	0.30	0	1
1995					
years of schooling	3457	12.00	2.87	7	19.5
Hauptschule	3455	0.40	0.49	0	1
Realschule	3455	0.34	0.47	0	1
Fachhochschulreife	3455	0.05	0.23	0	1
Abitur	3455	0.13	0.34	0	1
Apprenticeship	3457	0.69	0.46	0	1
University degree	3457	0.14	0.34	0	1

Source: GSOEP1985 and 1995 (100% version)

Table 2.5: Summary statistics on exogenous variables

Variable	N	Mean	Std.Dev	Min	Max
1985					
sex	4617	0.29	0.45	0	1
age	4617	37.18	10.06	20	55
experience	4617	20.70	10.54	0	43
tenure	4606	9.74	8.06	0	56.6
changed place since childhood	3181	0.60	0.49	0	1
1995					
sex	3457	0.31	0.46	0	1
age	3457	37.00	9.70	20	55
experience	3457	20.00	9.99	1	43
tenure	3457	9.74	8.77	0	41.3
changed place since childhood	2274	0.64	0.48	0	1

Source: GSOEP1985 and 1995 (100% version), own calculations

Table 2.6: Percentage of sample with given instrument status

individual grew up in ...		1985	1995
pc1	... a city	21.90	19.09
pc2	... a city or a big town	36.28	33.27
pc3	... some urban area	58.74	54.12

Source: GSOEP 1985 (N=4617) and 1995 (N=3457), own calculations

Table 2.7: OLS and IV results

	1985	1995
OLS		
	6.71	6.54
	(6.31;7.11)	(6.12;6.96)
IVE: place of childhood		
city (pc1)	12.63	12.58
	(7.89;17.39)	(8.45;16.70)
1st stage t	6.737	7.087
partial R ²	0.0098	0.0141
city or big town (pc2)	12.49	9.67
	(9.00;15.98)	(6.91;12.45)
1st stage t	9.247	9.691
partial R ²	0.0183	0.0265
urban (pc3)	13.28	9.22
	(7.94;18.63)	(6.95;11.48)
1st stage t	6.131	11.387
partial R ²	0.0081	0.0362
IVE: place of childhood * poor family background		
pc1*(poor fbq)	10.65	11.17
	(6.76;14.55)	(7.57;14.77)
1st stage t	-7.848	-7.361
partial R ²	0.0075	0.0105
pc2*(poor fbq)	9.86	11.29
	(7.44;12.28)	(8.34;14.25)
1st stage t	-11.721	-9.512
partial R ²	0.0142	0.0176
pc3*(poor fbq)	9.33	9.68
	(7.58;11.08)	(7.77;11.60)
1st stage t	-15.795	-13.845
partial R ²	0.0219	0.0273

Table 2.8: Actual average years of schooling by instrument status and family background quartile

	fbq1	fbq2	fbq3	fbq4
1985				
City	10.79	11.45	12.23	13.36
City or big town	10.66	11.36	12.19	13.54
Urban area	10.44	11.05	12.11	13.28
1995				
City	11.46	11.64	12.48	14.13
City or big town	11.59	11.57	12.49	13.97
Urban area	11.18	11.44	12.49	13.88

Source: GSOEP 1985 and 1995, own calculations

Table 2.9: Distribution of family background and individual variables across those 'counterfactual schooling quartiles'

Background quartile	1	2	3	4	Avg.
Father's education					
High school degree	0.00	0.26	2.43	20.85	5.76
Professional school	0.25	2.35	4.59	23.34	7.59
University degree	0.00	0.00	0.00	16.68	4.07
No schooling degree	35.30	23.67	0.09	0.09	15.01
Mother's education					
High school degree	0.00	0.00	0.35	6.83	1.75
Professional school	0.25	1.65	3.47	12.87	4.48
University degree	0.00	0.09	0.09	3.11	0.80
No schooling degree	42.29	30.72	0.87	0.80	18.93
Parental presence	0.00	30.64	96.01	95.21	54.86
Place of childhood					
City	19.63	20.19	22.36	25.55	21.90
City or big town	33.87	33.86	35.44	42.15	36.28
Urban	56.11	59.18	56.93	62.91	58.74
Change of place	55.46	61.74	62.02	58.86	60.14
Female	48.19	25.76	30.59	11.36	29.22
Mean age	40.59	35.90	38.04	34.00	37.18

frequency of respective characteristic by family background quartile; definition of quartiles based on regression of schooling level on family background variables (and age) for individuals from rural background and subsequent predictions for all observations as 'counterfactual schooling level if individual had grown up in a rural area'

Table 2.10: Covariate weights

	fbq1	fbq2	fbq3	fbq4
1985				
City	23.05	22.95	25.52	28.49
City or big town	24.00	23.22	24.42	28.36
Urban area	24.56	25.07	24.23	26.14
1995				
City	16.21	26.06	27.88	29.85
City or big town	18.00	25.91	26.61	29.48
Urban area	19.99	25.28	26.19	28.54

Note: w_q is the fraction in each quartile

Source: GSOEP 1985 and 1995, own calculations

Table 2.11: Differences in schooling by instrument status

	1985	1985	1995	1995
$Z = \dots$	0	1	0	1
City	11.32	12.04	11.85	12.59
City or big town	11.17	12.01	11.73	12.53
Urban area	11.10	11.74	11.56	12.36

Source: GSOEP 1985 and 1995, own calculations

Figure 2.1: Marginal benefit and marginal cost schedules for different individuals

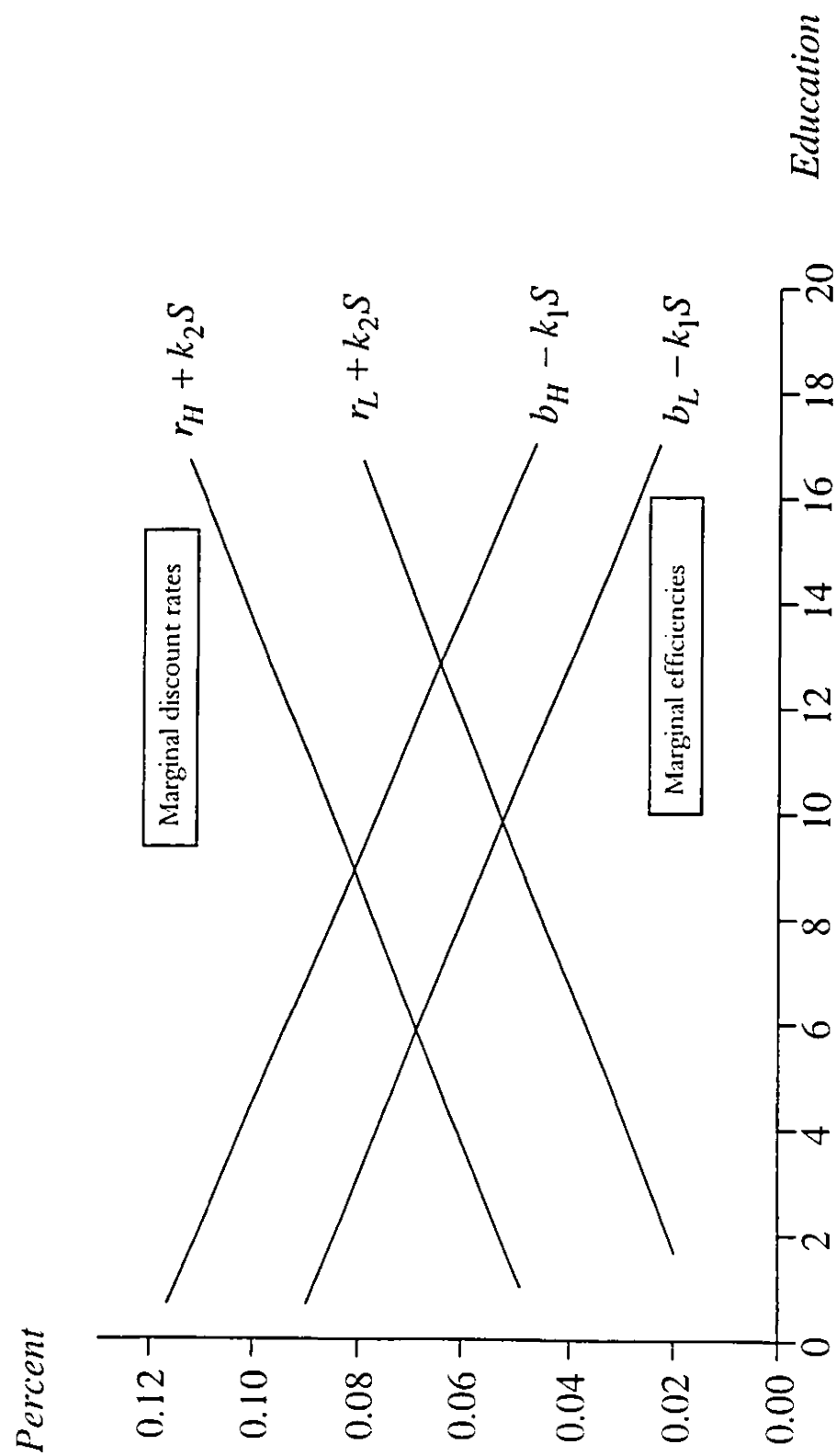
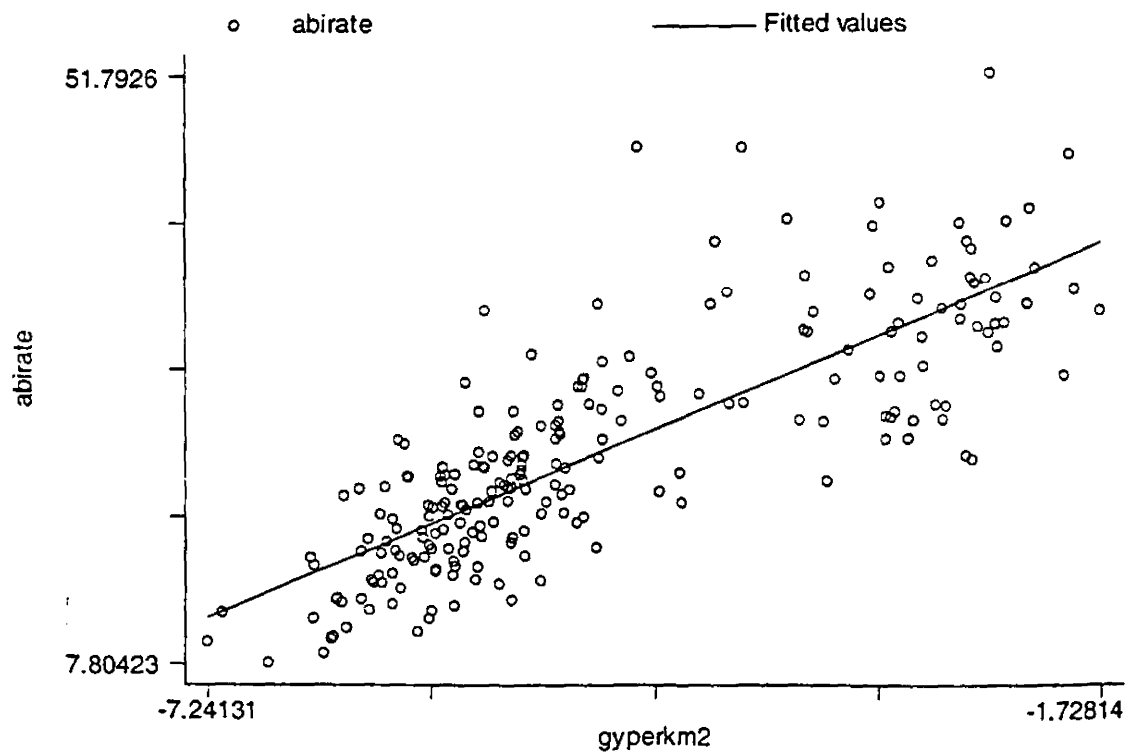


Figure 2.2: Educational attainment as a function of schooling infrastructure



Note: the figure plots the rate of school leavers having Abitur against the log of the number of "Gymnasium" per square kilometer

Source: data provided by the regional statistical offices (*Statistische Landesämter*) for the years 1996, 1997 or 1998 according to availability.

Figure 2.3: Actual average education difference by instrumental status using pc1
(1985 data)

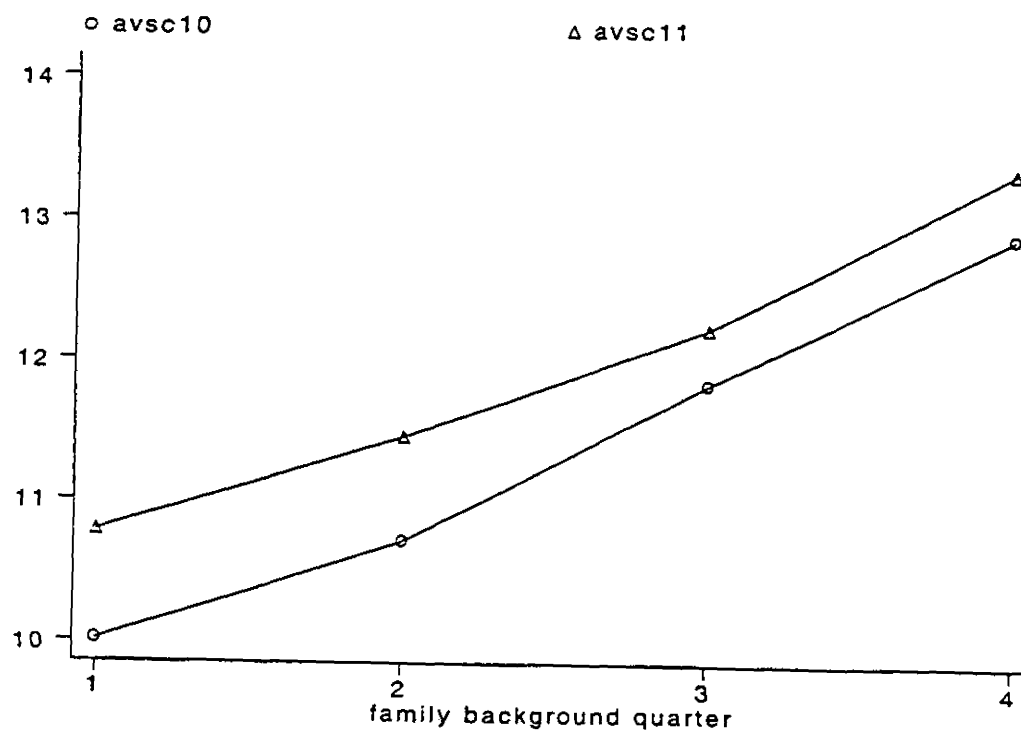


Figure 2.4: Actual average education difference by instrumental status using pc1
(1995 data)

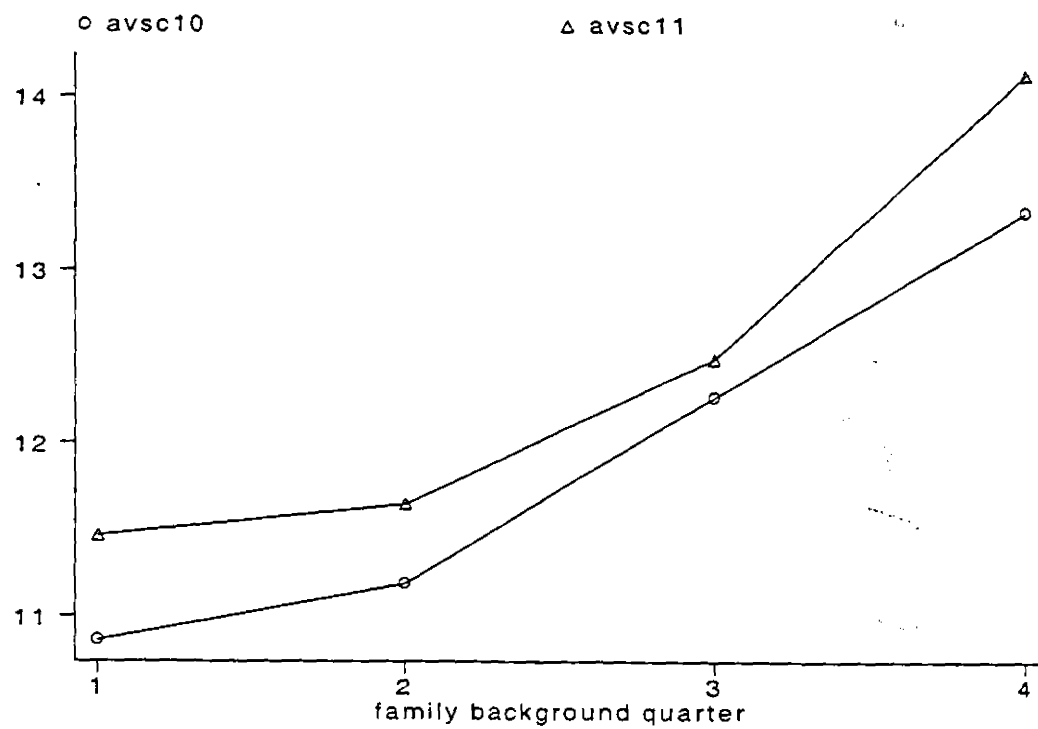
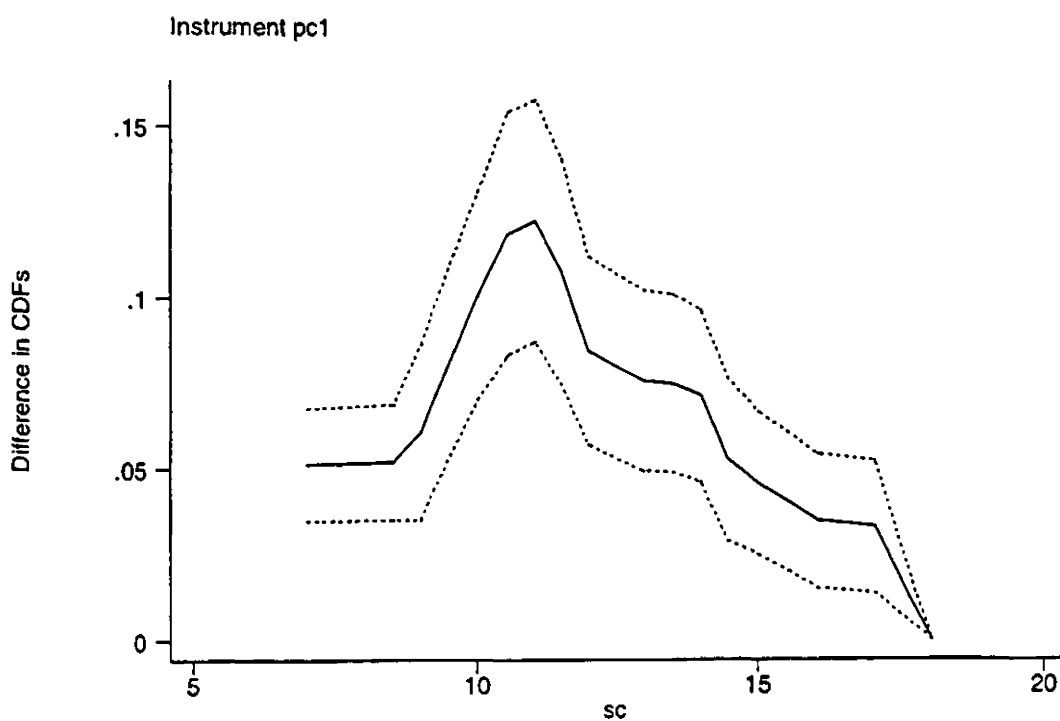


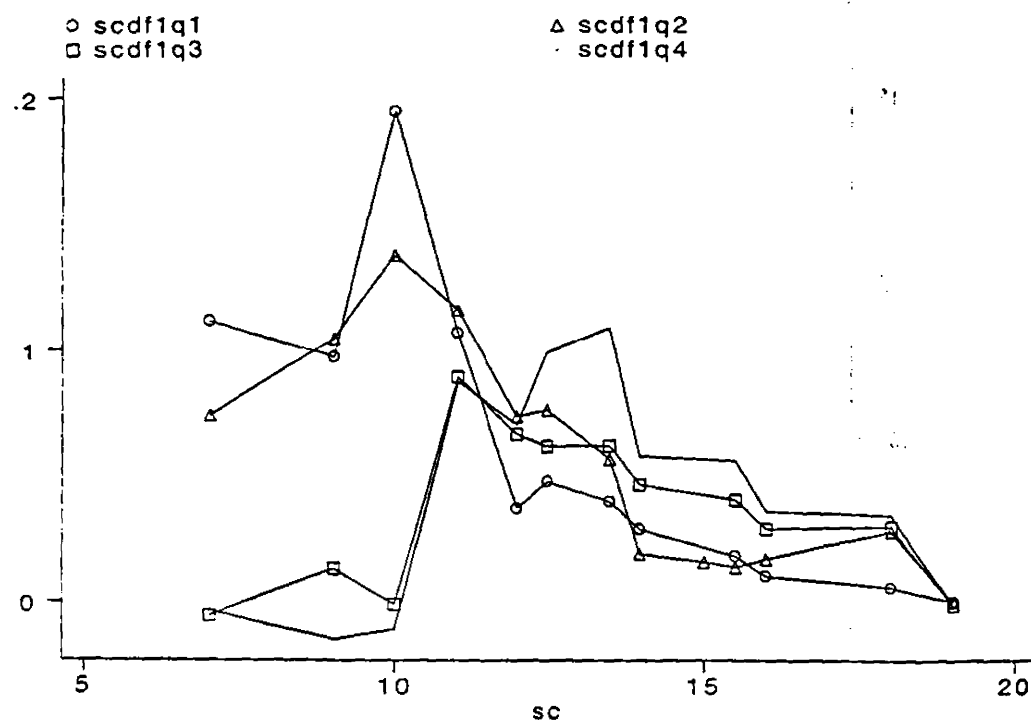
Figure 2.5: CDF difference using pc1 as an instrument (1985 data)

Note: the figure displays the fraction of the population who received at least one more year of schooling due to the instrument.

Calculated as the difference in the CDF: $\Pr(S < j | Z=0, X) - \Pr(S < j | Z=1, X)$

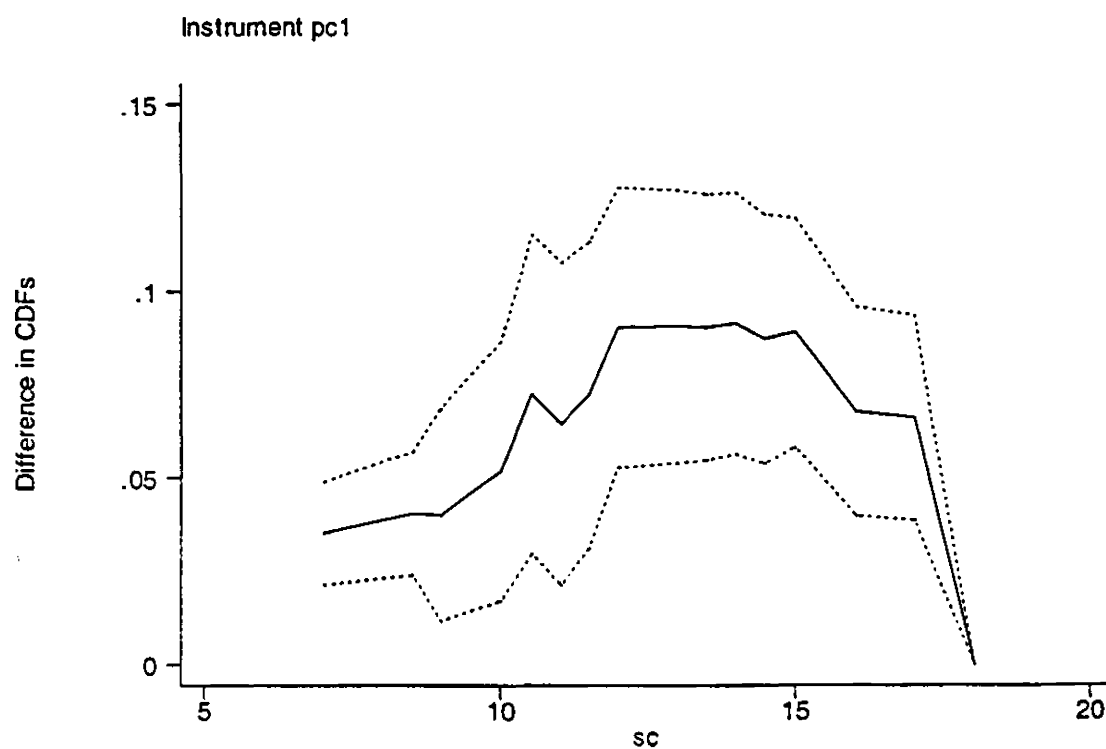
The 95% confidence bands are calculated using the conventional formula for a difference in proportions.

Figure 2.6: CDF difference by family background quartile using $pc1$ as an instrument (1985 data)



Note: the figure displays the fraction of the population who received at least one more year of schooling due to the instrument.

Calculated as the difference in the CDF: $\Pr(S < j | Z=0, X, Q) - \Pr(S < j | Z=1, X, Q)$

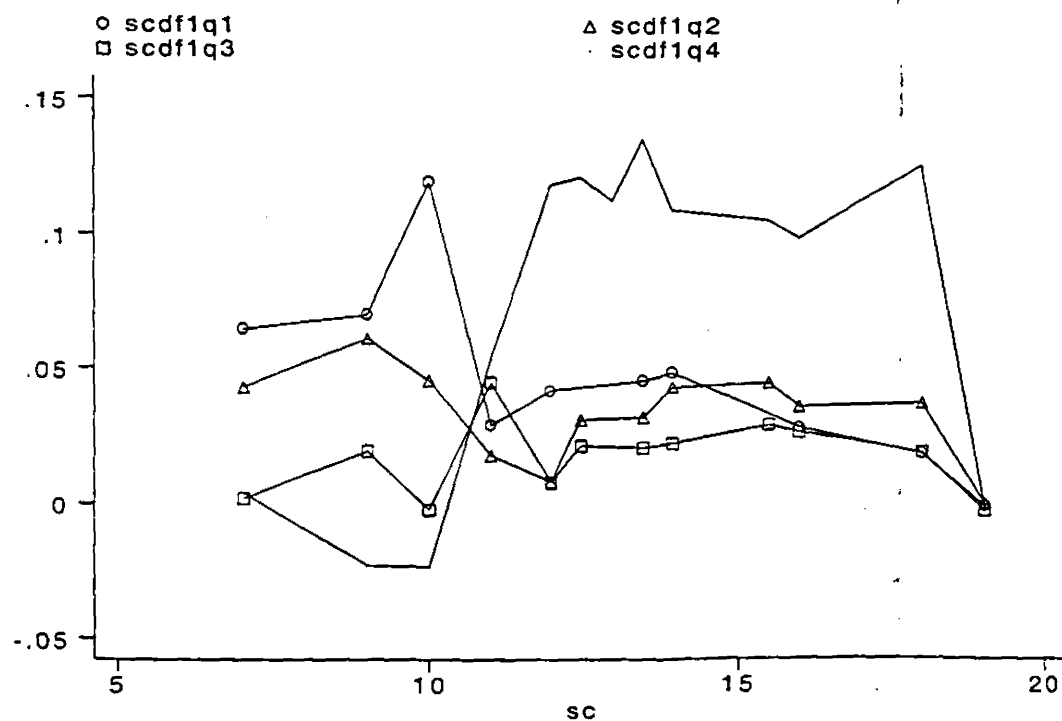
Figure 2.7: CDF difference using pc1 as an instrument (1995 data)

Note: the figure displays the fraction of the population who received at least one more year of schooling due to the instrument.

Calculated as the difference in the CDF: $\Pr(S < j | Z=0, X) - \Pr(S < j | Z=1, X)$

The 95% confidence bands are calculated using the conventional formula for a difference in proportions.

Figure 2.8: CDF difference by family background quartile using pc1 as an instrument (1995 data)



Note: the figure displays the fraction of the population who received at least one more year of schooling due to the instrument.

Calculated as the difference in the CDF: $\Pr(S < j | Z=0, X, Q) - \Pr(S < j | Z=1, X, Q)$

Bibliography

- [1] Angrist, Joshua D., 1990, "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records", *American Economic Review*, 80(3): 313-36.
- [2] Angrist, Joshua D. and Guido W. Imbens. "Two-Stage Least Squares Estimation of Average Causal Effects in Models With Variable Treatment Intensity", *Journal of the American Statistical Association*, Vol. 90, No. 430 (June 1995), 431-442.
- [3] Angrist, J.D. and A.B. Krueger (1991), *Does Compulsory School Attendance Affect Schooling and Earnings*, *Quarterly Journal of Economics* 106: 979-1014.
- [4] Angrist, J.D. and A.B. Krueger. "Empirical strategies in labor economics", Chapter 23 in: O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3, North-Holland (1999).
- [5] Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin. "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association*, Vol. 91, No. 434 (June 1996), 444-455.
- [6] Becker, Gary S. *Human Capital*, 1st ed. New York: Columbia University Press (for National Bureau of Economic Research), 1964.
- [7] Becker, Gary S. *Human Capital and the Personal Distribution of Income*, Ann Arbor: University of Michigan Press, 1967.
- [8] Bound, J., D.A. Jaeger and R.M. Baker (1995), Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak, *Journal of the American Statistical Association*, 90(430), 443-450
- [9] Boustedt, Olaf (1970): Pendlerverkehr. In: *Handwörterbuch der Raumforschung und Raumordnung*, 2. Aufl., Bd. II, Sp. 2284. Hannover.

- [10] Card, David. "Earnings, Schooling and Ability Revisited", in: Solomon Polachek, ed., *Research in Labor Economics*, Vol. 14 (1995a): 23-48.
- [11] Card, David, 1995b, Using Geographic Variation in College Proximity to Estimate the Return to Schooling, in Louis N. Christofides, E. Kenneth Grant, and Robert Swidinsky, eds., *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp* (University of Toronto Press, Toronto), 201-222.
- [12] Card, David. "The Causal Effect of Education on Earnings", in: O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3A (1999): 1801-63.
- [13] Heckman, James J., Lance Lochner, and Christopher Taber, "Human Capital Formation and General Equilibrium Treatment Effects: A Study of Tax and Tuition Policy", *Fiscal Studies* vol. 20, no.1 (1999): 25-40
- [14] Ichino, Andrea and Rudolf Winter-Ebmer. "Lower and Upper Bounds of Returns to Schooling: An Exercise in IV Estimation with Different Instruments," *European Economic Review* 43, no. 4-6 (1999) 889-901.
- [15] Ichino, Andrea and Rudolf Winter-Ebmer. "The Long-Run Educational Cost of World War Two", mimeo, 2000.
- [16] Imbens, Guido W. and Joshua D. Angrist. "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, Vol. 62, No. 2 (March, 1994), 467-475.
- [17] Kane, Thomas J. and Cecilia E. Rouse. "Labor Market Returns to Two- and Four-Year Colleges: Is a Credit a Credit and Do Degrees Matter?", NBER Working Paper No. W4268, January 1993.
- [18] Kling, Jeffrey, "Interpreting Instrumental Variables Estimates of the Returns to Schooling", *Journal of Business and Economic Statistics* (2001), forthcoming
- [19] Lang, K. (1993). "Ability Bias, Discount Rate Bias and the Returns to Education". Boston University, mimeo.
- [20] Lauer, Charlotte and Viktor Steiner. "Returns to Education in West Germany, An Empirical Assessment", Center for European Economic Research (ZEW), Mannheim, mimeo, 2000.

- [21] Lorenz, Wilhelm, and Joachim Wagner (1993), "A Note on Returns to Human Capital in the Eighties: Evidence from Twelve Countries", *Jahrbücher für Nationalökonomik und Statistik*, 211(1-2): 60-71.
- [22] Staiger, D. and J.H. Stock (1997), "Instrumental variables regression with weak instruments", *Econometrica*, 65, 557-586.
- [23] Wagner, Joachim, and Wilhelm Lorenz (1989), Einkommensfunktionsschätzungen mit Längsschnittdaten für vollzeiterwerbstätige deutsche Männer. (Estimates of Earnings Functions for Full-Time Working German Men Using Longitudinal Data), *Konjunkturpolitik*, 35(1-2): 99-109.

Chapter 3

Displacement in Germany

3.1 Introduction

Worker displacement, usually defined as the separation of workers for economic reasons, has attracted much attention in recent years. While it has been a subject of extensive studies in the US, interest in Europe has only recently arisen. Most US studies (e.g. Topel (1990) and Ruhm (1991)) find substantial and long-lasting effects of job loss on annual earnings and wages. In the short run, annual earnings of a typical blue-collar worker fall by 40 percent, most of which is caused by reduced labor supply (unemployment and weeks worked), though wage reductions are also substantial. In the longer run, workers' losses are mainly due to reduced wages, especially among experienced workers, those who are displaced from union jobs, and those who change occupation or industry after displacement.

Jacobson, LaLonde, and Sullivan (1993a and b) bring forward two additional results. First, displaced workers' losses begin mounting before separation. Second, losses are large even for those who find jobs in similar firms.

Stevens (1997) shows that much of the persistence in earnings losses can be explained by additional job losses.

Displacement is not only a negative personal experience but is also an important policy issue: where the short-run losses dominate, financial aid might be the major remedy. However, if the long-run losses are substantial, training programs giving workers the knowledge needed to return to higher-paid positions can be an important aid.

At the same time earnings losses following displacement can be informative about the empirical importance of models of wage formation like human capital theory, search and matching theory, and contract theory.

So, displacement is of both theoretical and policy interest and it is important to see whether the US findings translate to Europe and in particular Germany. Hitherto, there are only very few studies about displacement in Germany. Buttler and Bellmann (1991) use data from the German social insurance file (the so-called IAB¹ data set) to analyze earnings losses of displaced workers. Since the IAB data do not contain information on the cause of separations, they define displaced workers as workers who leave a job in an industry, where employment decreased between 1974 and 1986 by 30 percent or more. Their result is that earnings losses are restricted to older and unskilled workers while employees with a completed vocational training or professionals *gain* by leaving their firm.

Bender, Dustmann and Meghir (1998) use the IAB data together with an additional firm data set. Displacement is then defined as "losing one's job in a firm that closes within two years after separation". They state that displaced workers experience a positive wage growth throughout. Only when the sample is restricted to the group of displaced workers with an intervening unemployment spell of at least two years, the post-displacement wage is lower than the pre-displacement wage (the actual level depending on previous job tenure).

Burda and Mertens (2001), whose analysis is based on the same IAB data set, offer a different solution to the identification problem: starting with the GSOEP which contains information about the reasons for job displacement, they estimate a probit (1=displaced, 0=not displaced) on the sub-sample of the unemployed and use this to impute displacement in the IAB data set. They draw the conclusion that workers in the first pre-displacement wage quartile gain by displacement while the other three quartiles lose, the average wage loss being -3.64%.

Couch (2001) uses the GSOEP to study the effects of displacement due to firm closure. Since separate information on firm closure is only available from 1991 onwards, Couch restricts his sample to the period 1991-1996. He uses annual earnings and a count of annual months of unemployment as the dependent variables. Couch finds that in the year of displacement, annual earnings decline by about 13.5%, and that the typical worker experiences between 6 and 10 additional days of annual unemployment. Two years later, annual earnings are only 6.5% less than before displacement, and the largest estimated increase in annual unemployment is 4 days.

Grund (1999) uses the GSOEP to analyze possible stigma effects associated with dismissal as opposed to displacement due to plant closings.

This paper adds to the existing literature by analyzing heterogeneity in the costs of displacement between meaningful subgroups of workers. While Burda and Mertens consider

¹IAB stands for *Institut für Arbeitsmarkt- und Berufsforschung*, the institute that maintains the data.

differences in the costs of displacement across quartiles of the pre-displacement wage distribution, I construct a more comprehensive measure of the propensity of being displaced. In the spirit of propensity score matching, I estimate a probit regression of displacement in one of the years 1984-1995 on variables describing worker and firm characteristics in 1984. I use the predicted probability of displacement to regroup individuals into two meaningful subgroups. Comparing workers with equal probability of displacement but different actual treatment status, I find that workers *at high risk of being displaced* do not lose by displacement while those *at low risk of being displaced* lose about 16 percent in earnings in the years following displacement. Furthermore, I address a neglected issue in defining the sample period when using an econometric setup following Jacobson et al. My results overall corroborate the findings by Burda and Mertens.

The remainder of this paper is organized as follows: section 2 presents definitions of (the event of) displacement. Section 3 describes the data, gives an overview of the possibilities of defining the costs of displacement and relates them to the theoretical literature. Section 4 assesses the econometric methodology needed to analyze displacement. Section 5 presents the empirical results. Section 6 concludes and gives an outlook on possibilities for further work.

3.2 Defining Displacement

Theoretically speaking, displacement is defined as the involuntary termination of a position, excluding dismissals for cause.

In Germany, so far, mainly the social insurance file (IAB) has been used in empirical work. The main benefit of using an administrative data set like the IAB sample are its large sample size and its (supposedly) higher reliability. These advantages, however, come at a cost. The IAB data only report the event of job separation while the GSOEP also gives the (self-reported) reason for job separation (quit, layoff, maternity leave etc.). Furthermore, the GSOEP has a richer set of demographic variables.

Given the lack of information on the reason of job separation, studies using the IAB data have to define or impute displacement in some way. Bellmann and Butler (1991) define displaced workers as workers who leave a job in an industry, where employment decreased between 1974 and 1986 by 30 percent or more. Their definition necessarily misses displaced workers in industries that do not contract by as much as 30 percent. Bender, Dustmann, and Meghir (1998) define as displaced those workers who separate from firms which close within two years before closure. Their group of displaced therefore does not include laid off people

in ongoing firms. Burda and Mertens (2001) offer a different solution to the identification problem: starting with the GSOEP which contains information about the reasons for job displacement, they estimate a probit (1=displaced, 0=not displaced) on the sub-sample of the unemployed and use this to impute displacement in the IAB data set. As they acknowledge, this introduces measurement error because "[s]ome individuals will be predicted as displaced when not (false positives or Type I error) while others are classified as not displaced when in fact they are (false negatives or Type II error)."

In contrast, workers' responses to survey questions allow us to identify all layoffs (though probably including layoffs for cause). However, responses might be less precise than firms' records. Survey questions usually only ask for the last job change and we might therefore miss multiple job changes. In addition, there is usually a considerable amount of recall bias in timing the date of layoff.

In any case it is interesting and important to investigate displacement with both kinds of data sets and this paper tries to contribute to this issue by exploiting the GSOEP information.

3.3 The data

3.3.1 The German Socioeconomic Panel (GSOEP)

For the analysis of displacement, the GSOEP (see Holst, Lillard, and DiPrete, 2001, and the appendix for further details) contains all the information necessary. In the annual surveys, individuals are asked about their current labor market status and retrospective information going back to the beginning of the previous calendar year is collected. In case of a job change since the previous survey, respondents are asked the reason for job change: "Why did you leave this job? Which one of the following applies to you?" and can choose among several answers including *terminated by employer*, *firm closed down* (from 1991 on) and *quit the job*. I define individuals as displaced if they declare *terminated by employer* or *firm closed down* as reasons for a job change since the previous survey.

Moreover, the GSOEP contains a rich set of job-related and demographic characteristics that I use as control variables in the estimation. The variables used reflect factors known to influence labor market behavior and outcomes: age and age squared, a set of dummy variables indicating education and vocational training (less than high school degree, high school degree or equivalent, more than high school degree), firm size, industry and job tenure. Table 3.1 shows some descriptive statistics referring to the base year 1984 for the sample used in the

estimations.

I exclude from the analysis civil servants, self-employed, trainees, and workers in agriculture. I drop observations with missing information on one of the variables. Furthermore, I drop full-time employed individuals with unreasonably low monthly gross incomes of less than 1000 DM.² The whole analysis below is based on samples A (West-Germans) and B (foreigners) of the GSOEP.³ I use a dummy for foreigners. In view of the fixed effects regressions below, I only keep individuals with at least two years of non-missing observations.

Before discussing the dependent variables, I show some descriptive statistics about the distribution of job changes.

3.3.2 Descriptive statistics

Table 3.2 is based on waves 2 through 12 (year 1985 through 1995) of the GSOEP and captures all job changes (i.e. job-to-job shifts, job-to-unemployment shifts, and job-to-nonemployment shifts).

The percentage of the displaced⁴ among all job separators - those whose employment relation ends - is 15.4 percent. More revealing for the importance of job displacement is the next column: 36.8 percent of those who experienced a job change since the last interview and are still unemployed at the current interview have been displaced. The above table shows another fact which is in accordance with theory: The percentage of job quitters (voluntary leave) among all job separations is 32.5 percent while it is considerably lower (15.9 percent) for those still unemployed at the interview date. Quitters find a job much more easily (and probably only quit if they have already found a new job ?!) than the displaced.

Table 3.3 shows the percentage of displaced to non-displaced job separators across occupations. Since some occupations only contain very few observations I had to aggregate some groups.

Considering the whole period, 1985-1995, there is a remarkable difference in occupational distribution between displaced and non-displaced job separators. Production workers are more likely to be displaced than to separate for other reasons.

²Results are robust with respect to this cut-off point.

³Note that the foreign sample consists of households where the household head is of Turkish, Greek, Yugoslavian, Spanish, or Italian nationality (the five largest groups of foreign nationals). It consists mainly of 'guest workers' who came to Germany in the 1950's and 1960's already and who have therefore already assimilated to the native German population.

⁴Those people answering that the reason for job change was "termination by the employer" or "company closed down".

A more comprehensive picture of the factors characterizing displacement is provided by the probit estimation beneath.

3.3.3 The outcomes

This section describes the two measures of the cost of displacement used in this study and gives a brief overview of related theoretical models.

The costs of displacement are typically measured in terms of either wages or labor earnings. These variables are informative about different dimensions of the costs of displacement.

1) Labor earnings are the most direct measure of the cost of displacement and capture the monetary loss associated with displacement. This measure contains the effect of both wage changes and in labor supply. Simply put, earning 50% less after displacement can be due to both earning half of the previous wage or earning the same wage as before but working half-time. It is therefore a relatively atheoretic, descriptive measure of the costs of displacement.

2) The variable of interest in theoretical models, however, usually is the wage rate because it relates to marginal productivity, wage premia, and wage seniority. I use gross monthly income in the month before the interview. Since I restrict the sample to the full-time employed, using monthly income is equivalent to using hourly wages. I choose gross income because this measures the firm's valuation of the worker as opposed to net income which makes results less clear by mixing them up with tax effects. The progressive German tax system compresses the wage structure in a way that provides insurance (subsidizes) against the cost of displacement because lower marginal tax rates in lower earnings categories is equivalent to a smaller net displacement loss as opposed to the gross displacement cost.

The most prominent explanation of wage premia and wage seniority is human capital theory (Becker, 1975). The main distinction is between general human capital and sector- and/or firm-specific human capital. General human capital cannot explain earnings losses following displacement since it is transferable across firms. However, human capital accumulated in a specific firm (sector) might be of limited use in another firm (sector). Therefore, the return to investment in specific human capital given by a higher wage for the employee can be destroyed upon displacement. In the new position, the worker starts with his stock of general human capital but with no specific capital. Consequently, there will be no wage premium in the new job.

Search and matching theory is another candidate to explain earnings losses. Pioneered by Jovanovic (1979a and b), the basic idea of this literature is that both firms and workers are heterogeneous and therefore good matches have to be searched for. In the sequel,

highly-productive matches will survive longer than bad matches and will be more highly compensated. In case the match is destroyed, earnings losses will occur on the new job. Put differently, to find a highly-paid reemployment possibility workers have to search longer.

Lazear's model (1981) of a promotion-from-within policy yields another explanation of wage seniority and corresponding particularly high earnings losses for high-tenure workers: to give young workers the incentive to stay with the firm, they are paid a low wage (below marginal productivity) with the prospect of earning a high wage (above their marginal productivity) when older. Within this model, losing a job means beginning in jobs farther down the promotion ladder.

This short overview of theoretical models motivates why we might expect income losses following displacement but I do not attempt to test those theories.

3.4 The econometric approach

A simple approach to measuring earnings losses is to take the difference between post- and pre-displacement earnings as dependent variable and regress it on a number of explanatory variables. This measure has several shortcomings: it only captures the short-run effects of displacement and it does not take into account the (counterfactual) earnings growth that might have occurred in the absence of displacement.

The following more sophisticated approach takes potential earnings into account: denote by D_{is} the binomial variable that takes the value 1 if worker i was displaced at date s , and 0 otherwise. Then, we can compare actual earnings at date t following displacement to the expected earnings had the worker not been displaced:

$$E(y_{it}|D_{is} = 1) - E(y_{it}|D_{is} = 0) \quad (3.1)$$

Since this does not rule out that worker i was displaced at some other time than date s , which would be the case of multiple displacements, it makes sense to focus on the case where there is the single displacement date s :

$$E(y_{it}|D_{is} = 1) - E(y_{it}|D_{ik} = 0 \text{ for all } k) \quad (3.2)$$

This latter expression better captures the effect of job loss on workers' earnings and career than *one* job loss out of several consecutive ones.⁵

⁵The costs associated with multiple job losses are a different issue which has been analyzed by Stevens (1997).

In a next step, we will want to condition on worker characteristics in addition to displacement status. One way to do this is to condition on I_{is} , the information set containing information about worker i at the date of displacement:

$$E(y_{it}|D_{is} = 1, I_{is}) - E(y_{it}|D_{ik} = 0 \text{ for all } k, I_{is}) \quad (3.3)$$

According to Jacobson et al. (1993a), this captures only part of the adverse effects associated with displacement. Since firm conditions leading to displacement might already affect earnings in the periods before this event, a full image of displacement should include these adverse pre-displacement losses. This effect is illustrated in Figure 3.1.

Assume that p periods ($p = 4$ in Figure 3.1) before displacement, it becomes known that displacement is going to take place at date s . Then, earnings of displaced workers can be compared to earnings that these workers expected to receive at date $s - p$:

$$E(y_{it}|D_{is} = 1, I_{is-p}) - E(y_{it}|D_{ik} = 0 \text{ for all } k, I_{is-p}) \quad (3.4)$$

The information set should of course include variables that influence earnings.

The econometric implementation of (3.4) can be achieved by estimating the following augmented version of a standard earnings equations à la Mincer (1974):

$$y_{it} = \alpha_i + \gamma_t + x_{it}\beta + \sum [d_in4, \dots, d_0, \dots, d4ago]_{it} * \delta + \varepsilon_{it} \quad (3.5)$$

which includes a set of dummy variables representing the event of displacement. The notation is as follows: $d_ink_{it} = 1$ in year t if worker i is going to be displaced in year $t + k$, $d_0_{it} = 1$ if he is going to be displaced in this year⁶, and $d_kago_{it} = 1$ if worker i was displaced in year $t - k$. This scheme is explained in the following simple example for a worker who was displaced in 1990:

	d_in4	d_in3	d_in2	d_in1	d_0	$d1ago$	$d2ago$	$d3ago$	$d4ago$
1988	0	0	1	0	0	0	0	0	0
1989	0	0	0	1	0	0	0	0	0
1990	0	0	0	0	1	0	0	0	0
1991	0	0	0	0	0	1	0	0	0

Note that we cannot use all years of data (1984-1995) in the regressions because the displacement dummies are only well-defined in the central period of data (1988-1991 in our

⁶More than 90% of the interviews takes place in February, March, and April.

case where the dummy window stretches out from $t - 4$ to $t + 4$).⁷ In the years outside the central period, we have no information about treatment status. In the first year of data, 1984, for example, we do not know if a worker was displaced in one of the previous years, and therefore the dummies $d4ago$, $d3ago$, $d2ago$, and $d1ago$ are not well-defined. By the same token, in the last year of data, 1995, we do not know if a worker is going to be displaced in one of the subsequent years, and therefore the dummies d_in1 , d_in2 , d_in3 , and d_in4 are not well-defined. This is a crucial point often neglected by other studies. In their seminal paper Jacobson et al.(1993) use all years of data in their regressions and set those dummies equal to zero and thereby introduce a potential bias by misclassifying some (actually) treated units as controls. A correct treatment of the issue requires to first choose the set of dummies thought to describe the displacement event and to adjust the period of data used in the regressions accordingly.

The remaining variables are included to control for other general factors that influence earnings. The vector x_{it} includes observed, time-varying characteristics of the worker. γ_t captures the general time pattern of earnings in the economy and α_i measures the worker-specific effect. The error term ε_{it} is assumed to be of constant variance and to be uncorrelated across individuals and time. The control group is the group of workers that are continuously employed throughout the whole sample period.⁸

3.4.1 Treatment of the unemployed

The treatment of displaced workers unemployed in the month before the interview and therefore not reporting monthly income needs special discussion. Jacobson et al. set their quarterly income to zero and not e.g. to the level of unemployment benefits. This choice is arbitrary and partly explains the huge losses of up to 40% reported for US workers. I decided to omit the relevant person-year observation from the wage regression. This might introduce sample selection if those unemployed after displacement are not a random subgroup of the displaced. Sample selection models in panel data are computationally very demanding (see Kyriazidou, 1997). If the sample size was bigger, I could estimate a selection (into unemployment) equation on the sample of the displaced only and then introduce a selection term in the earnings equation. Taking into consideration that the sample size is small, it is not possible to follow this approach. For this reason, the results should be interpreted

⁷1986-1993 in the case where the dummy window stretches out from $t - 2$ to $t + 2$. Other cases follow accordingly.

⁸More precisely, our comparison group are all workers that are employed at all interviews, and have never been displaced.

conditional on employment.

3.5 Empirical Analysis

3.5.1 OLS results

We start out by showing results of the OLS regression⁹

$$y_{it} = \gamma_t + x_{it}\beta + \sum [d_in4, \dots, d_0, \dots, d4ago]_{it} * \delta + \varepsilon_{it} \quad (3.6)$$

i.e. (3.5) ignoring the worker-specific fixed effect α_i . We use monthly gross earnings as dependent variable. Remember that we use income data for the period 1988-1991 whereas data for all the years 1984-1995 are utilized to define the displacement dummies. The results shown in Table 3.4 give a purely descriptive picture of the earnings differences between displaced and non-displaced workers and should not be interpreted causally.

In addition to the control variables described above, I include time dummies to control for business cycle effects.

Turning to the coefficients of the displacement dummies it appears that the displaced already earn about 4 to 5 percent less¹⁰ than the control group before displacement though the effect is not statistically significant. There is a significant drop in earnings in the year immediately before displacement and earnings differences of up to 15 percent in the years after displacement. The coefficient in the year after displacement (*d1ago*) which is smaller (and statistically insignificant) than the other post-displacement coefficients might be explained by the fact that it only captures those workers that immediately find a (comparable) job while those unemployed at the interview have a missing observation in that year and only reenter estimation in the second year after displacement.

Hitherto, I have not included job tenure as an explanatory factor. This is done in Table 3.5 by introducing a dummy for job tenure of more than 4 years.¹¹ Pre-displacement earnings differences are unaffected and post-displacement earnings differences become smaller but remain significant. Thus, displacement losses can not only be explained by a loss in seniority.

Now, I want to go beyond the basically descriptive picture given by the OLS results which mixes up effects *due to* displacement and permanent differences between workers.

⁹Note that we employ OLS with the cluster option which adjusts standard errors for the fact that observations of the same individual i in different years are not independent of each other.

¹⁰Remember that in semilogarithmic regressions, the percentage change induced by a dummy variable with coefficient δ switching from 0 to 1 is $e^\delta - 1$ (Halvorsen and Palmquist (1980)).

¹¹The results do not change when using a quadratic in tenure.

That means that we return to (3.5) which explicitly takes into account individual-specific fixed effects.

3.5.2 Fixed Effects Regression I

Here, I present results of fixed effects regressions based on the same sample of workers used in the OLS regression.

Table 3.6 shows that, conditioning on individual-specific effects, we only see a significantly negative effect of displacement on wages in the year of (i.e. immediately before) displacement. In addition, there is a significantly (though small in absolute value) positive effect on wages four years after displacement. This suggests that - overall - displacement is not associated with any substantive losses.

However, it could be the case that we measure this insignificant overall effect because the implicit assumption of constant coefficients on the displacement dummies across subpopulations does not hold. As a consequence, different and significant effects on subgroups might cancel out on average and therefore yield the overall insignificant effect of displacement. To give an example, men's and women's earnings might react differently to displacement and simply controlling for gender will only affect the constant term in the regression but not the slopes. The same might be true for differences between native Germans and foreigners or for differences across workers' industries etc.

The most clear-cut way to proceed would be to divide the whole sample into cells defined by worker and job characteristics. We could then compare earnings patterns of non-displaced workers and displaced workers that are "identical" to each other. This (statistical) matching is already prevented by data limitations because we would end up with too few workers per subgroup. Alternatively, we could divide the whole sample only according to one dimension (e.g. sex) at a time. Yet, this approach would produce a whole set of singular results. For this reason, I try to find a composite measure encompassing all the factors that are relevant for displacement.

3.5.3 Displacement probit

We can make this idea operational by estimating a probit in 1984 - i.e. at the beginning of the whole sample period - predicting the probability that a worker with certain personal and job characteristics will be displaced in one of the years until 1995. I will classify a worker to be *at high risk of being displaced* if his predicted probability of being displaced exceeds the average predicted displacement probability. Similarly, a worker is *at low risk of*

being displaced if his predicted probability of being displaced lies below this threshold.¹² In this way we homogenize subgroups of workers along several dimensions instead of dividing workers by some arbitrary (single) variable. The idea behind this first step is to create meaningful subgroups of workers that might have different earnings patterns (i.e. different coefficients on the displacement dummies) due to displacement. This approach is reminiscent of the propensity score approach because it uses pre-treatment variables to estimate the probability of treatment (displacement) and then compares outcomes (earnings patterns) of individuals with similar characteristics, i.e. having a similar propensity score.¹³

In addition, it is quite intuitive to consider these subgroups: given their overall characteristics, some workers are more likely to be displaced and others are less likely to be displaced. Still, in both subgroups we have some workers that are actually displaced and some that remain employed as can be seen from Table 3.8 based on the results from the probit estimation:

The probit is of course also of interest in itself. Table 3.7 shows that it is a mix of characteristics that determines displacement. The displaced are on average younger, earn less, and have lower tenure on the job. Workers in small firms (with less than 20 workers) are more likely to be displaced. Firing is easier for these firms because German legislation only requires works councils for firms with more than 20 workers.

One surprising finding is that there is no effect of gender on the probability of being displaced. Interestingly enough, there are also no significant differences across education levels.

3.5.4 Fixed Effects Regression II

We can now turn to the results of the fixed-effects estimation on the two subgroups: They are different from the US findings (Jacobson et al. (1993)). There, earnings start decreasing already before displacement which can be interpreted as the worker's firm being in bad condition and therefore cutting wages (or weekly hours) before finally displacing the worker. In Germany, these wage cuts do not show up - to a great extent because of German labor

¹²As shown by Cramer (1999), it is a general feature of predicted values in probit estimation based on unbalanced samples (i.e. with unequal sample frequencies) that "the less frequent outcome always has lower estimated prediction probabilities than the other". The within-sample percentage correctly predicted - taking 0.5 as a cut-off level - as implemented in many statistical packages would therefore classify too many people as being at low risk of being displaced.

¹³In a follow-up paper, I try to implement propensity score estimation of the effect of treatment (displacement) on the treated (the displaced). See the last section for an outlook on this project.

market regulations. The pre-displacement behavior of wages is therefore easy to understand.

More interesting is the difference between the two subgroups of *high risk and low risk workers* (tables 3.9 and 3.10). The fact that earnings of the *high risk* displaced recover so quickly can be seen as an indication that these workers find comparable jobs after displacement. We saw in the probit estimation that these are mainly workers who had lower earnings anyway which cannot drop much further because collective wage agreements imply de facto minimum wages in most industries. In contrast, those *at low risk of being displaced* used to earn more to begin with and had higher job tenure. As a separate analysis on only the continuously employed workers shows, tenure is a significant factor explaining earnings for the subgroup of *low risk workers* while it has no explanatory power for those *at high risk of being displaced*. This finding corresponds with the intuition that those *at low risk of being displaced* have profited from sizeable rents which they lose upon displacement. These results are in line with dual labor market theory (see Saint-Paul, 1996). In the primary layer of the labor market, jobs are quite safe and workers earn more than workers in the secondary labor market. As the results show, higher job security in the primary labor market comes at a price: when being laid off, these workers lose more than those laid off from the secondary labor market.

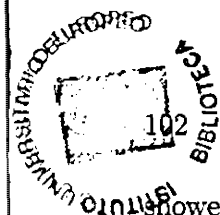
The results presented so far were all based on monthly gross earnings as dependent variable. I repeated the above estimation with yearly labor income. The results are very similar as regards the general pattern of the displacement coefficients. However, now the main effect occurs in the year of displacement (d_0), the coefficient being -0.21 for the *high risk* workers and -0.26 for the *low risk* workers. This indicates that on average the displaced lose a lot in yearly labor income because many of them go through a period of unemployment.

To sum up, it is unemployment following displacement which hurts a considerable share of the displaced in the short run, while only those at low risk of being displaced have a long-run negative effect on earnings.¹⁴

3.6 Conclusion and outlook

In this paper I presented theoretical definitions of displacement and described ways to identify them in empirical work. Then, I suggested variables measuring the cost of displacement and

¹⁴I also estimated a Cox proportional hazard model to explain survival times in unemployment. Displaced workers going through unemployment do not differ from all other unemployed suggesting that there is no particular stigma effect for the displaced. This result ties in with Grund (1999) who does not find any stigma effects associated with dismissal as opposed to displacement due to (supposedly exogenous) plant closings.



CHAPTER 3. DISPLACEMENT IN GERMANY

showed to which theoretical models they are related.

The empirical part showed important differences with US findings: contrary to the US, in Germany earnings do not start decreasing before displacement. Post-displacement wage losses do not occur for all workers. I defined and identified two subgroups of workers: those *at high risk of being displaced* and those *at low risk of being displaced*. Within the first subgroup, those actually displaced do not incur earnings losses compared to their non-displaced counterparts. In contrast, the displaced in the second group lose up to 16% in earnings in the second and third year after displacement, and therefore have to bear a long-lasting cost of displacement.

My results are in line with the results by Burda and Mertens who also find heterogeneity in wage losses. They estimate wage losses to be most pronounced for workers in the upper three pre-displacement wage quartiles. To the extent that the pre-displacement wage distribution is a main correlating factor of the propensity of being displaced, my results yield a robustness check for their finding. Furthermore, my results go beyond theirs by showing that losses are not restricted to the short run.

The approach of re-grouping workers into two more homogeneous groups, defined by similar propensity of being displaced, goes some way towards comparing likes and thereby towards establishing a clean causal effect of displacement on earnings. Yet, we can improve on the current procedure by rigorously employing propensity score matching. The propensity score is defined by Rosenbaum and Rubin (1983) as the conditional probability of receiving the treatment given the pre-treatment variables. The probit estimation above was essentially a propensity score estimation. However, the current approach falls short of establishing the *balancing property* of the propensity score¹⁵ according to which observations with the same propensity score must have the same distribution of observable covariates independently of treatment status.

In a follow-up paper, I use the algorithm outlined in the appendix to estimate the propensity score. The average treatment effect on the treated (ATT) is typically estimated using one of several estimators proposed in the evaluation literature, e.g. the *Stratification* method, the *Best Matching* method and the *Radius Matching* method (see Heckman et al. 1998 and Lechner 1999). It is not straightforward to apply one of these estimators to this case. My treatment is non-standard because I do not have a single treatment dummy but treatment is described by the whole set of dummies ($d_{in4}, \dots, d_{0}, \dots, d_{ago}$). For this reason, I cannot simply compute the ATT as a single number but the ATT is given by a whole set of numbers (namely the coefficients of the displacement dummies in a regression). The idea

¹⁵See again Rosenbaum and Rubin (1983).

of the follow-up paper is to implement an equivalent of the *Stratification* method in the following way: by regressing - block by block¹⁶ - log income on a constant term and the set of displacement dummies. Since the treatment spans an extended period of time (recall that treatment is defined as displacement in any of the years 1984-1995) and some observable characteristics that affect earnings might change independent from displacement, I include contemporary variables as further controls in the regressions. To control for unobservable factors in earnings potentials, fixed effects estimators are used to purge individual-specific fixed effects. The displacement losses for the overall population are given by the weighted average of displacement losses across blocks where the weights are given by the number of treated observations in the block. It is of course informative by itself to see if and how displacement losses vary by block. Alternatively, I could run a single regression with a full set of displacement-dummy*block-interactions.¹⁷

The main problem in applying the propensity score matching procedure to the GSOEP data is probably the relatively small sample size. Still, it seems to be a promising approach to treat displacement along the lines of the program evaluation literature.

¹⁶Using the blocks generated by the algorithm to estimate the propensity score.

¹⁷Both approaches are reminiscent of the "conditional difference-in-differences" approach outlined by Heckman, Ichimura, Smith, and Todd (1998). They propose to apply difference-in-differences - and fixed effects is similar - to a sample, that is already matched on pre-treatment observable variables, in order to remove also unobservable differences.

3.7 Appendix

3.7.1 Algorithm for the estimation of the propensity score

The algorithm to estimate the propensity score consists of the following steps:

1. Estimate the probit model:

$$Pr\{D_i = 1|X_i\} = \Phi(h(X_i))$$

where Φ denotes the normal c.d.f. and $h(X_i)$ is a starting specification which includes all the covariates as linear terms without interactions or higher order terms.

2. Split the sample in 5 equally spaced intervals of the propensity score.
3. Within each interval test that the average propensity score of treated and control units do not differ.
4. If the test fails in one interval, split the interval in halves and test again.
5. Continue until, in all blocks, the average propensity score of treated and control units do not differ.
6. Within each interval, test the balancing property, i.e. that the means of each covariate do not differ between treated and control units.
7. If the means of one or more covariates differ, start again from the beginning with a different specification of $h(X_i)$.

3.7.2 Figures and Tables

Figure 3.1: possible earnings pattern

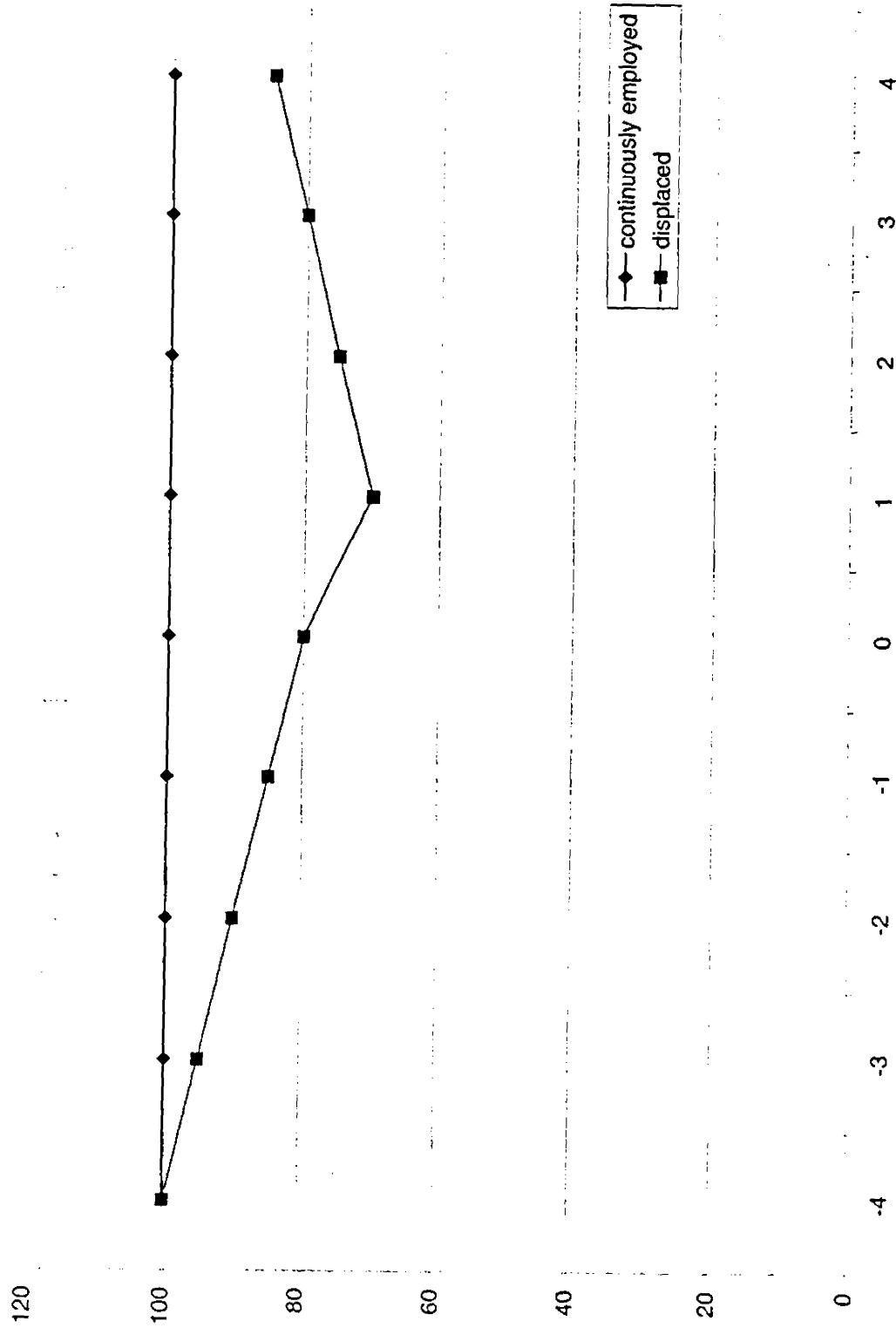


Table 3.1: Descriptive statistics for the year 1984

Variable	Mean	Std.Dev.
Individual characteristics		
Female	0.20	0.4
Foreign	0.34	0.47
Age	37.31	8.29
Age ² /100	14.61	6.09
Less than high school	0.45	0.48
High school or equiv.	0.48	0.5
More than high school	0.07	0.26
Job/firm characteristics		
Log(gross monthly inc.)	7.8	0.31
Long tenure (>4 years)	0.82	0.38
Small firm (<20)	0.13	0.34
Industry dummies		
Energy	0.01	0.12
Mining	0.01	0.08
Manufacturing	0.58	0.49
Construction	0.11	0.32
Trade	0.05	0.23
Transport	0.04	0.19
Banking, Insurance	0.04	0.19
Services	0.15	0.36
Number of observations	861	

Source: Own calculations based on GSOEP, wave 1, subsamples A & B

Table 3.2: Type of jobs separations

reason for job shift	all job separations	workers unemployed at interview date
displaced	761	328
quit	1605	142
end of fixed term contract	589	155
other	1989	267
total number of observ.	4944	892

Source: Own calculations based on GSOEP, waves 1-12, subsamples A & B

Table 3.3: Distribution of the displaced across occupations

Occupation	1985-1995		1985-1990		1991-1995	
	Displ.	Not Displ.	Displ.	Not Displ.	Displ.	Not Displ.
Execut.&Profess.&Technic.	11.5	18.8	8.6	19.1	14.0	18.3
Sales&Adminis.Support	29.1	37.3	25.8	37.4	31.9	37.2
Service Jobs	8.1	8.9	9.3	7.2	7.0	10.9
Production Jobs	48.8	32.4	53.6	33.3	44.7	31.3
Other	2.5	2.6	2.7	2.9	2.3	2.2

Source: Own calculations based on GSOEP, waves 1-12, subsamples A & B

Table 3.4: OLS Regression with robust standard errors, 1988-1991

Variable	Coefficient	Std.Error
Age	0.041	0.009
Age ² /100	-0.046	0.011
1989	0.010	0.005
1990	0.035	0.006
1991	0.047	0.007
Less than high school	-0.022	0.024
More than high school	0.457	0.048
Foreign	-0.189	0.023
Female	-0.237	0.019
Displacement in 4 years	-0.042	0.026
Displacement in 3 years	-0.018	0.027
Displacement in 2 years	-0.043	0.028
Displacement in 1 years	-0.048	0.031
Displacement this year	-0.079	0.029
Displacement 1 year ago	-0.054	0.052
Displacement 2 years ago	-0.148	0.068
Displacement 3 years ago	-0.143	0.034
Displacement 4 years ago	-0.088	0.042
Constant	7.147	0.181
Number of obs = 3288		
Number of persons = 861		
R-squared = 0.4034		

Notes: Dependent variable: log of gross monthly income. Further controls: industry dummies.

Source: GSOEP, subsamples A & B

Table 3.5: OLS Regression with robust standard errors controlling for tenure, 1988-1991

Variable	Coefficient	Std.Error
Age	0.040	0.009
Age ² /100	-0.045	0.011
1989	0.009	0.005
1990	0.036	0.006
1991	0.047	0.007
Less than high school	-0.023	0.024
More than high school	0.460	0.049
Foreign	-0.189	0.023
Female	-0.239	0.019
Tenure>4	0.055	0.036
Displacement in 4 years	-0.038	0.026
Displacement in 3 years	-0.016	0.027
Displacement in 2 years	-0.036	0.028
Displacement in 1 years	-0.044	0.031
Displacement this year	-0.073	0.029
Displacement 1 year ago	-0.008	0.061
Displacement 2 years ago	-0.108	0.075
Displacement 3 years ago	-0.101	0.046
Displacement 4 years ago	-0.058	0.046
Constant	7.104	0.184
Number of obs = 3288		
Number of persons = 861		
R-squared = 0.4046		

Notes: Dependent variable: log of gross monthly income. Further controls: industry dummies.

Source: GSOEP, subsamples A & B

Table 3.6: Fixed effects regression on the whole sample, 1988-1991

Variable	Coefficient	Std.Error
Age	0.065	0.009
Age ² /100	-0.053	0.010
1989	-0.006	0.005
1990	0.002	0.005
Tenure>4	0.016	0.016
Displacement in 4 years	-0.013	0.014
Displacement in 3 years	-0.009	0.016
Displacement in 2 years	-0.021	0.018
Displacement in 1 years	-0.017	0.021
Displacement this year	-0.048	0.022
Displacement 1 year ago	-0.004	0.031
Displacement 2 years ago	0.006	0.029
Displacement 3 years ago	0.012	0.024
Displacement 4 years ago	0.043	0.021
Constant	6.164	0.198
Number of obs = 3288		
Number of persons = 861		
R-sq: within = 0.0671		

Notes: Dependent variable: log of gross monthly income.

Source: GSOEP, subsamples A & B

Table 3.7: Displacement Probit estimates 1984

Variable	Coefficient	Std.Error
Log(gross monthly inc.)	-0.125	0.061
Age	-0.041	0.015
Age ² /100	0.058	0.020
Tenure>4*	-0.166	0.046
Small firm (<20)*	0.123	0.050
Less than high school*	-0.061	0.040
More than high school*	0.018	0.066
Female*	0.003	0.042
Foreign*	0.087	0.044

(*) dF/dx is for discrete change of dummy variable from 0 to 1

Observed probability 0.235
 Predicted probability 0.217 (at x-bar)

Number of observations = 861

LR chi2(12) = 78.25 Prob > chi2 = 0.0000

Log likelihood = -429.937

Pseudo R2 = 0.0834

Notes: Dependent variable equal to 1 if person displaced in one of the years 1984-1995, 0 otherwise.

Further controls: industry dummies.

Reference groups: service sector, German, men, high school degree, big firm

Source: GSOEP, subsamples A & B

Table 3.8: Predictions based on Probit estimates 1984

	predicted	displaced	
actually displaced	0	1	total
0	461	198	659
1	81	121	202
total	542	319	861

Table 3.9: Fixed effects regression on the subsample of those at high risk of being displaced, 1988-1991

Variable	Coefficient	Std.Error
Age	0.076	0.014
Age ² /100	-0.062	0.016
1989	-0.013	0.009
1990	0.002	0.009
Tenure>4	0.024	0.024
Displacement in 4 years	-0.025	0.022
Displacement in 3 years	-0.011	0.024
Displacement in 2 years	-0.004	0.027
Displacement in 1 years	0.004	0.033
Displacement this year	-0.054	0.034
Displacement 1 year ago	0.024	0.044
Displacement 2 years ago	0.070	0.040
Displacement 3 years ago	0.062	0.033
Displacement 4 years ago	0.080	0.028
Constant	5.683	0.327

Number of obs = 1196

Number of persons = 319

R-sq: within = 0.0841

Notes: Dependent variable: log of gross monthly income.

Source: GSOEP, subsamples A & B

Table 3.10: Fixed effects regression on the subsample of those at low risk of being displaced, 1988-1991

Variable	Coefficient	Std.Error
Age	0.059	0.012
Age ² /100	-0.048	0.014
1989	-0.001	0.005
1990	0.004	0.005
Tenure>4	-0.012	0.022
Displacement in 4 years	0.003	0.019
Displacement in 3 years	-0.008	0.022
Displacement in 2 years	-0.043	0.025
Displacement in 1 years	-0.053	0.029
Displacement this year	-0.037	0.031
Displacement 1 year ago	-0.094	0.047
Displacement 2 years ago	-0.166	0.047
Displacement 3 years ago	-0.134	0.043
Displacement 4 years ago	-0.076	0.038
Constant	6.432	0.264
Number of obs = 2092		
Number of persons = 542		
R-sq: within = 0.0774		

Notes: Dependent variable: log of gross monthly income.

Source: GSOEP, subsamples A & B

Bibliography

- [1] Angrist, Joshua D. and Alan B. Krueger. "Empirical Strategies in Labor Economics." in: Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, vol.3. Amsterdam: North-Holland, 1999.
- [2] Ashenfelter, Orley A. and David E. Card. "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs." *Review of Economics and Statistics* 67, no.4 (1985) 648-60.
- [3] Becker, Gary S. *Human Capital*, 2nd ed. New York: Columbia University Press (for National Bureau of Economic Research), 1975.
- [4] Bender, Stefan, Christian Dustmann, and Costas Meghir. "Job Displacement, Non-employment and Wages in Germany." *mimeo*, 1998.
- [5] Bender, Stefan, Jürgen Hilzendegen, Götz Rohwer, and Helmut Rudolph. "Die IAB-Beschäftigtenstichprobe 1975-1990." *Beiträge zur Arbeitsmarkt- und Berufsforschung* 197, 1996.
- [6] Büchtemann, C.F. "Kündigungsschutz als Beschäftigungshemmnis ?" *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 23 (1990): 394-410.
- [7] Burda Michael C. and Antje Mertens. "Estimating Wage Losses of Displaced Workers in Germany", *Labour Economics* (8)1 (2001) pp. 15-41.
- [8] Buttler, Friedrich, and Lutz Bellmann. "Mikroanalyse des Arbeitsmarktes mit Individualdaten der Sozialversicherungsträger: Einkommensverluste bei Freisetzen." *Ifo-Studien* 37, no.3-4 (1991) 297-314.
- [9] Couch, Kenneth A. "Earnings Losses and Unemployment of Displaced Workers in Germany" *Industrial and Labor Relations Review* 54, no.3 (2001): 559-572.

- [10] Cramer, J.S. "Predictive Performance of the Binary Logit Model in Unbalanced Samples." *Journal of the Royal Statistical Society: Series D (The Statistician)* 48, no.1 (1999): 85-94.
- [11] Gerlach, Knut, and Ulrich Schasse. "Arbeitsmarktwirkungen von Kündigungen und Entlassungen." In *Betriebszugehörigkeitsdauer und Mobilität - Theoretische und Empirische Analysen*, Gerlach, Knut and Olaf Hübler, eds., Arbeitspapier 1990-4 aus dem SAMF (1990): 121-42.
- [12] Grund, Christian. "Stigma Effects of Layoffs? Evidence from German micro-data", *Economics Letters* 64 (1999): 241-247.
- [13] Hall, Robert E. "Lost Jobs." *Brookings Papers on Economic Activity* (1995): 221-56.
- [14] Halvorsen, Robert and Raymond Palmquist. "The Interpretation of Dummy Variables in Semilogarithmic Equations." *American Economic Review* 70, no.3 (1980): 474-75.
- [15] Heckman, J., H. Ichimura, J. Smith, and P. Todd "Characterizing Selection Bias Using Experimental Data", *Econometrica*, 66 (1998), 1017-1098.
- [16] Jacobson, Louis S., Robert LaLonde, and Daniel Sullivan. *The Costs of Worker Dislocation*. Kalamazoo, MI: W.E. Upjohn Institute, 1993a.
- [17] Jacobson, Louis S., Robert LaLonde, and Daniel Sullivan. "Earnings Losses of Displaced Workers," *American Economic Review*, 83, no.4 (1993b): 685-709.
- [18] Jovanovic, Boyan. "Job Matching and the Theory of Turnover." *Journal of Political Economy* 87, no.5 (1979a): 972-90.
- [19] Jovanovic, Boyan. "Firm-Specific Capital and Turnover." *Journal of Political Economy* 87, no.6 (1979b): 1246-60.
- [20] Kletzer, Lori G. "Job Displacement" *Journal of Economic Perspectives* 12, no.1 (1998): 115-36.
- [21] Lazear, Edward. "Agency, Earnings Profiles, Productivity, and Hours Restrictions." *American Economic Review* 71, no.4 (1981): 606-20.
- [22] Lechner, Michael. "Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption", in: Lechner, M., Pfeiffer, F.

- (eds), *Econometric Evaluation of Labour Market Policies*, Heidelberg: Physica/Springer (2001) pp. 43-58.
- [23] Rosenbaum, P.R., and D.B. Rubin "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika* 70, no.1 (1983), 41-55.
- [24] Ruhm, Christopher J. "Are Workers Permanently Scarred by Job Displacements?" *American Economic Review* 81, no.1 (1991): 319-324.
- [25] Saint-Paul. Gilles. *Dual labor markets: a macroeconomic perspective*. Cambridge: MIT Press, 1996.
- [26] SOEP Group. "The German Socio-Economic Panel (GSOEP) after more than 15 years - Overview." In: Elke Holst, Dean R. Lillard und Thomas A. DiPrete (eds.): *Proceedings of the 2000 Fourth International Conference of German Socio-Economic Panel Study Users (GSOEP2000)*. *Vierteljahrshefte zur Wirtschaftsforschung* 70, no.1 (2001): 7-14.
- [27] Stevens. Ann Huff. "Persistent Effects of Job Displacement: The Importance of Multiple Job Losses." *Journal of Labor Economics*, 15, no.1 (1997): 165-188.
- [28] Topel. Richard. "Specific Capital and Unemployment: Measuring the Costs and Consequences of Job Loss," *Carnegie-Rochester Conference Series on Public Policy* 33 (1990): 181-214.

Chapter 4

Risk-Sharing in the Short Run and In the Long Run

(Written with Mathias Hoffmann)

4.1 Introduction

Do industrialised countries use the same channels to insure against long-term and short-term income risks? Do they insure in different ways against different types of shocks? This paper aims to provide an answer to these questions.

The starting point of our analysis is the observation that most countries' consumption risks do not seem to be internationally diversified. French and Poterba (1991) were the first to hint at the huge home biases in international equity portfolios. This non-diversification puzzle has been cast into various formulations that are not only based on stocks of foreign assets but also on flow variables. Most notably, Backus, Kehoe and Kydland (1992) demonstrated that international consumption correlations are too low to be explained by models with perfect capital mobility and complete asset markets.

A complementary perspective on international non-diversification is provided in a series of papers by Asdrubali, Sørensen and Yosha (1996) and Sørensen and Yosha (1998). Asdrubali, Sørensen and Yosha (1996), henceforth ASY, suggested a simple decomposition of income risk that allows the investigator to distinguish between the cross-sectional and the intertemporal dimension of risk sharing. The cross-sectional dimension is reflected in the cross-border ownership of state-contingent assets such as equity or in fiscal transfer schemes. The intertemporal dimension is reflected in borrowing and lending, i.e. in the use of national

or international credit markets.

In US state-level data ASY (1996) find that 39 percent of shocks to gross state product are smoothed by capital markets, 13 percent are smoothed by the federal government, and 23 percent are smoothed by credit markets. Conversely, Sørensen and Yosha (1998) find that EU and OECD countries achieve much less cross-sectional risk sharing than do US states. Mélitz and Zumer (1999) extended the ASY study by including further exogenous variables like regional size and the real interest rate. Their results by and large corroborate those of ASY.

The findings of the line of research surveyed here - and to which this paper aims to contribute - may have far-reaching implications for the prospective workings of the European Monetary Union. A plethora of studies documents that, in terms of Mundell's (1961) classical criteria, Europe is much less of an optimum currency area than are e.g. the United States. In particular, macroeconomic shocks are generally found to be much less symmetric in Euroland than among US states. Against this background, finding out to what degree capital markets can contribute to the insurance of aggregate income and consumption risk has become a question of paramount importance: if shocks are asymmetric, perhaps they can be smoothed through sufficient risk sharing. A monetary union that experiences asymmetric macroeconomic disturbances may not appear optimal when measured against the classical OCA criteria but it may provide a huge pool of risks that can be optimally insured - as long as the channels mentioned above are actually available and do get exploited.

Sørensen and Yosha (1998) conclude quite negatively in this respect: given that European countries do not seem to exploit risk sharing opportunities, EMU could entail high welfare costs in the absence of intensified fiscal transfer mechanisms. Mélitz and Zumer on the other hand conclude that the start of monetary union will promote the sharing of risks via market channels.

In this paper, we extend the method of ASY (1996) to a fully dynamic framework¹. In so doing, we use recently developed methods for the estimation of panel vector autoregressions. Our method allows us to assess how income uncertainty at short and long forecast horizons is insured. It also allows us to investigate how different types of income uncertainty get insured. The most important distinction to be made along these lines is the one between permanent and transitory shocks to income. Insurance against permanent idiosyncratic shocks requires perpetual claims on some sort of income that is negatively correlated with a country's own income stream. Conversely, transitory fluctuations can be completely smoothed through

¹ Athanasoulis and van Wincoop (2000) provide an analysis of risk-sharing at various time horizons but their model does not allow the identification of risk-sharing channels.

borrowing and lending. *A priori*, we should therefore expect that permanent shocks get insured through different channels than transitory shocks and our econometric model allows us to disentangle these two types of shocks with minimal identifying restrictions.

Our results can be summarized as follows:

- Short-term and long-term risks are insured in the same ways. The forecast horizon does not matter for either the choice of insurance channel nor for the extent to which income risk is insured overall.
- Insurance against transitory shocks to income is generally much better than against permanent shocks and is achieved largely through credit markets, i.e. the intertemporal risk sharing channel. This result ties in with the theoretical findings by Baxter and Crucini (1995) who demonstrated that, as long as shocks are not too persistent, the full risk sharing allocations that pertain in models with complete asset markets can be very well approximated in models that only feature non-state contingent borrowing and lending. Whereas Baxter's and Crucini's work provides a theoretical rationalization of our results, a recent empirical study by Kraay, Loayza, Serven and Ventura (2000) has demonstrated that countries' international portfolios are largely held in the form of international credit rather than equity. This is the empirical corollary that may explain the importance of the credit channel for the sharing of transitory income risks.
- Earlier results in the literature suggested that capital markets provide only a minimal share of the total consumption insurance that is achieved between countries. Even though the role of capital markets for consumption insurance remains limited once we condition on the type of shock, their role seems much more respectable than would appear from our unconditional dynamic setup or from the results obtained in Sørensen and Yosha (1998).
- There is some evidence of insurance of permanent shocks through the intertemporal channel. The reason for this could be that a big share of a country's GDP cannot be traded on capital markets, e.g. because labour income is non-insurable. This may give rise to precautionary savings. Athanasoulis and Shiller (2000) have shown how the degree of market incompleteness affects the incidence of precautionary savings. Our results lend further empirical support to their view.
- Overall, roughly 60 percent of income variability in industrialised countries remains uninsured, most of it due to a failure to insure against permanent fluctuations in income.

Our results may also have important implications for further research into the sources of the home bias. It is generally found that national capital markets do much better in providing insurance to regions than do international capital markets in providing insurance to countries (compare for example the results in Asdrubali, Sorensen and Yosha (1996) and Sorensen and Yosha (1998)). Also, international asymmetries in output fluctuations are generally much more persistent than intranational ones (see for example Chamie et al. (1994)). At the same time, our results reveal that the failure of international capital markets to provide insurance is particularly due to a lack of insurance of 'permanent' income risks. An important question that future research should address is therefore why international capital markets do so badly in providing insurance against permanent shocks.

The remainder of the paper is organized as follows:

Section two outlines our dynamic econometric model of international risk sharing. In our empirical implementation, we rely on a panel vector autoregression that we implement using the method suggested by Holtz-Eakin, Newey and Rosen (1988). Our approach has the advantage that we only have to rely on cross-sectional asymptotics. Furthermore, we can identify permanent and transitory shocks to output with only minimal identifying assumptions by exploiting equilibrium relations between the data. Section three describes the data and our empirical results. We report the results of our analysis in section four and we offer conclusions in section five.

4.2 A dynamic model of risk sharing

In this section we propose a dynamic econometric model that enables us to analyse how income risk is shared over time. More specifically, our model allows us to identify the relative roles of intertemporal (i.e. borrowing and lending) and cross-sectional smoothing (i.e. insurance through international capital markets)².

The starting point of our analysis is the following decomposition of the variance of per capita GDP-growth:

$$\begin{aligned} \text{var}(\Delta gdp_t | \mathcal{I}_{t-1}) &= \text{cov}(\Delta gdp - \Delta gnp, \Delta gdp_t | \mathcal{I}_{t-1}) \\ &\quad + \text{cov}(\Delta gnp - \Delta c, \Delta gdp_t | \mathcal{I}_{t-1}) \\ &\quad + \text{cov}(\Delta c, \Delta gdp_t | \mathcal{I}_{t-1}) \end{aligned} \tag{4.1}$$

²Our method is closely related to Asdrubali, Sorensen and Yosha (1996). Their approach is completely static, however. Another difference between our model and that of ASY is that we do not allow for a fiscal insurance channel. Sorensen and Yosha (1998) have demonstrated that the fiscal channel is not important for the international dimension of risk sharing which is what we focus on in this paper.

Lower case letters denote logarithms and gnp and c denote gross national product and consumption per capita respectively. The conditioning information set \mathcal{I}_{t-1} contains realizations of variables that are known at the end of period $t - 1$.

We can divide (4.1) by $var(\Delta gdp_t | \mathcal{I}_{t-1})$ to get:

$$1 = \beta_C + \beta_S + \beta_U$$

where

$$\beta = \begin{bmatrix} \beta_C \\ \beta_S \\ \beta_U \end{bmatrix} = \frac{1}{var(\Delta gdp_t | \mathcal{I}_{t-1})} \begin{bmatrix} cov(\Delta gdp - \Delta gnp, \Delta gdp) \\ cov(\Delta gnp - \Delta c, \Delta gdp) \\ cov(\Delta c, \Delta gdp) \end{bmatrix} | \mathcal{I}_{t-1} \quad (4.2)$$

The individual coefficients in β can now be associated with various channels of risk sharing. The GDP-GNP differential reflects international factor income flows. Hence, β_C measures to what extent capital income from abroad covaries with GDP. Therefore, β_C can be thought of as representing the cross-sectional dimension of consumption insurance that is achieved (primarily) through cross-border ownership of foreign assets, i.e. through international capital markets.³

The GNP-C differential measures savings and β_S gives the contribution of the intertemporal aspect of consumption insurance (i.e. smoothing through savings). Finally, β_U is the residual covariance between consumption growth and GDP growth, reflecting the undiversified or unsmoothed component of consumption.

We will now describe how we identify the conditional variances and covariances involved in (4.1). For this purpose, let

$$\Delta X_t = \begin{bmatrix} \Delta gdp_t \\ \Delta gdp_t - \Delta gnp_t \\ \Delta gnp_t - \Delta c_t \end{bmatrix}$$

Then we assume that

$$\mathcal{I}_t = \{X_\tau\}_{\tau=1}^t \quad (4.3)$$

and that expectations coincide with linear projections. These assumptions allow us to express $E(\Delta X_t | \mathcal{I}_{t-1})$ as a vector autoregression. The unexpected component of ΔX_t which we will denote by ε_t is now given by the reduced-form residual of the VAR:

$$\Phi(L)\Delta X_t = \varepsilon_t \quad (4.4)$$

³As Sorensen and Yosha (1998) note, labour income flows between industrialised countries are negligible. The same holds true for interest payments on international bonds and loans. We can therefore think of the GDP-GNP differential as a good proxy for contingent capital income such as equity returns.

where $\Phi(L)$ is a 3×3 matrix polynomial in the lag operator, L , which satisfies the condition that the roots of $\det(\Phi(z))$ lie outside the unit circle.

Now let Ω denote the variance-covariance matrix of ε_t and let ω_{ij} be the entry in the i -th row and j -th column of Ω . Then

$$\beta_c = \frac{\omega_{21}}{\omega_{11}} \text{ and } \beta_s = \frac{\omega_{31}}{\omega_{11}}$$

Of course, the analogue of β_u is given by

$$\beta_u = 1 - \beta_c - \beta_s \quad (4.5)$$

We can now generalize our approach to arbitrary forecast horizons in order to answer the question as to what the role of various channels for risk sharing at these horizons may be. The mean squared prediction error in a VAR, k periods ahead, is given by

$$\Psi(k) = MSPE_k = \sum_{i=0}^{k-1} C_i \Omega C_i' \quad (4.6)$$

where the C_i are the matrix coefficients of the moving average representation of ΔX_t .

Let the entries of $\Psi(k)$ be denoted by $\psi_{ij}(k)$. Then the analogue of β from above can be defined:

$$\beta_c(k) = \frac{\psi_{21}(k)}{\psi_{11}(k)} \text{ and } \beta_s(k) = \frac{\psi_{31}(k)}{\psi_{11}(k)}$$

and again

$$\beta_u(k) = 1 - \beta_c(k) - \beta_s(k)$$

and

$$\beta(k) = \begin{bmatrix} \beta_c(k) & \beta_s(k) & \beta_u(k) \end{bmatrix}'$$

Obviously, $\beta(1) = \beta$ because $C_0 = I$ and therefore $\Psi(1) = \Omega$.

Note also that as the forecast horizon gets infinite, $\beta(k)$ should converge to the unconditional β that emerges from the static ASY model. Hence, the basic ASY regression provides a check of specification for any VAR estimation that may provide the basis for the dynamic decomposition given in (4.6). We are now going to deal with estimation issues.

4.2.1 PVAR Estimation

A naive application to an individual country of the procedure outlined in the previous section, is not likely to yield meaningful results. Estimating separate VARs for each country would not allow us to control for country fixed effects, possibly leading to seriously biased estimates.

Also, we need to take into account time varying fixed effects that are common to a whole cross-section of countries. This is because common or global shocks cannot be insured and we have to make sure that they do not pollute our estimates. We will therefore employ panel techniques in order to identify country-specific and time-specific components by exploiting not only the time series but also the cross-sectional dimension. As our sample period is relatively short - we employ annual data from 1975 to 1990 - we will rely on dynamic panel methods that are robust to short samples, i.e. require only cross-sectional asymptotics.

To see the problems that are associated with estimating this type of dynamic panel model, write out the standard reduced-form representation (4.4) to get

$$\Delta X_t = \mu + \sum_{l=1}^p \Phi_l \Delta X_{t-l} + \varepsilon_t \quad t = p+1, \dots, T \quad (4.7)$$

Reinterpreting this as a system of panel equations yields

$$\Delta X_{it} = \mu + \sum_{l=1}^p \Phi_l \Delta X_{i,t-l} + \lambda_t + f_i + u_{it} \quad i = 1, \dots, K; t = p+1, \dots, T \quad (4.8)$$

where now all variables vary by i and t , and where f_i is the vector of country-specific effects and λ_t is a time-specific effect. Since ΔX_{it} is a function of f_i , $\Delta X_{i,t-1}$ is also a function of f_i . Therefore, $\Delta X_{i,t-1}$, a right-hand regressor in (4.8), is correlated with the error term. This renders the OLS estimator biased and inconsistent even if the u_{it} are not serially correlated. For the standard fixed effects (FE) estimator, the 'within' transformation wipes out the country-specific effects f_i , but $(\Delta X_{i,t-1} - \bar{\Delta X}_{i,-1})$ where $\bar{\Delta X}_{i,-1} = \sum_{t=2}^T \Delta X_{i,t-1} / (T-1)$ will still be correlated with $(u_{i,t} - \bar{u}_{i,-})$ even if the u_{it} are not serially correlated. This is because $\bar{u}_{i,-}$ contains $u_{i,t-1}$ which is correlated with $\Delta X_{i,t-1}$ by construction. In the technical appendix, we describe how we have used instrumental variables techniques following the method set out by Holtz-Eakin, Newey and Rosen (1988) to estimate the model given in (4.8).

In the remainder of this section, we are going to discuss how we can incorporate permanent and transitory shocks in the VAR-model (4.8). Since in what follows, panel notation will generally not be required, we will henceforth drop the index i or the fixed effects in our discussion.

4.2.2 Permanent shocks and risk sharing

Our interest in this paper is in the comovement of *growth rates* of consumption and various output aggregates. Still, it is possible that the *levels* of these variables may have feedback

effects on the growth rates. To the extent that output and consumption are likely to follow integrated processes, such feedbacks from the level-variables would imply cointegrating relations between the variables.

In standard models of economic growth cointegrating relationships will most likely arise in the form of a stationary 'great ratio' of consumption over output (see for example King et al. (1991) and Neusser (1991)). In our setup, the great ratio is just given by the difference $c - gnp^4$. A formal test based on the dynamic panel OLS procedure suggested by Nelson and Sul (2001) strongly rejected the null of no cointegration between idiosyncratic GNP and consumption. We therefore decided to add this variable as an error correction term to the panel VAR model given in (4.14). Our model accordingly looks as follows:

$$\Delta X_{it} = \Phi \Delta X_{it-1} + \gamma \delta' X_{it-1} + \varepsilon_{it} \quad (4.9)$$

where $\delta' = \begin{bmatrix} 0 & 0 & -1 \end{bmatrix}$ is the cointegrating vector and $\gamma = \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix}'$ represents the vector of adjustment coefficients.

An error-correction model such as (4.9) allows the identification of permanent and transitory disturbances without further identifying assumptions. Following Johansen (1995), the permanent shocks can be written as

$$\pi_{it} = \gamma'_{\perp} \varepsilon_{it} \quad (4.10)$$

whereas the transitory disturbances are identified by requiring that they be orthogonal to the space of permanent shocks. Hence

$$\tau_{it} = \gamma' \Omega^{-1} \varepsilon_{it} \quad (4.11)$$

Note that whenever π_t or τ_t are non-scalar, the permanent or transitory shocks are not identified among themselves. However, for our purposes, this does not matter. The share of the forecast error variance that is explained by all permanent or transitory shocks does not depend on how we identify each of these shocks individually.

To see this assume that S_{π} and S_{τ} are appropriately dimensioned non-singular matrices such that $\pi_0 = S_{\pi} \pi$ and $\tau_0 = S_{\tau} \tau$. Let furthermore, as in the non-cointegrated case, $C(L)$

⁴We measure the great ratio as C/GNP , not, as is common, as C/GDP . The reason for this is, that at least in principle, a country's GDP and consumption can diverge arbitrarily if foreigners own perpetual claims on a sufficiently large share of that country's income. This is exactly what we should see if risk sharing was perfect.

be the reduced form matrix polynomial giving the Wold-representation of (4.9). Then the structural, i.e. just-identified form is given by $C(L)P^{-1}$ where

$$P = \begin{bmatrix} S_{\pi}\gamma'_{\perp} \\ S_{\tau}\gamma'\Omega^{-1} \end{bmatrix}$$

is just the matrix mapping the reduced-form disturbances into their permanent and transitory components.

It is easily verified that

$$P^{-1} = \begin{bmatrix} \Omega\gamma_{\perp}(\gamma'_{\perp}\Omega\gamma_{\perp})^{-1}S_{\pi}^{-1} & \gamma(\gamma'\Omega^{-1}\gamma)^{-1}S_{\tau}^{-1} \end{bmatrix}$$

Then note that the covariance of $\begin{bmatrix} \pi_0 & \tau_0 \end{bmatrix}'$ is given by

$$\Sigma = \begin{bmatrix} S_{\pi}\gamma'_{\perp}\Omega\gamma_{\perp}S'_{\pi} & 0 \\ 0 & S_{\tau}\gamma'\Omega^{-1}\gamma S_{\tau} \end{bmatrix}$$

Hence, the mean-square prediction error is

$$\Psi(k) = \sum_{l=0}^{k-1} C_l \left[\Omega\gamma_{\perp}(\gamma'_{\perp}\Omega\gamma_{\perp})^{-1}\gamma'_{\perp}\Omega + \gamma(\gamma'\Omega^{-1}\gamma)^{-1}\gamma' \right] C'_l \quad (4.12)$$

where the first term in parentheses measures the contribution of permanent shocks and the second the transitory. It can be seen from (4.12) that $\Psi(k)$ is independent of any particular choice of S_{π} and S_{τ} . Hence, the relative contributions of permanent and transitory shocks do not depend on the particular just-identification chosen.

We are going to report the estimation results for the cointegrated panel VAR and the ensuing decomposition of the prediction error in section 4.

4.3 Data and Empirical Implementation

We used annual per capita data for GDP, GNP and consumption (C), for 23 industrial countries, from the Penn World Tables (PWT, release 5.6). We generated world per capita aggregates of each of the three variables using population data from the same source. Annual observations on all three variables were available for the period 1970-90. In our estimation, we included only the period 1975-90 in order to avoid potential parameter instability in the model that is bound to arise if the oil shock and the aftermath of the demise of Bretton-Woods was included. Following Sørensen and Yosha (1998), we did not extend the sample

beyond 1990 to avoid instability problems that are likely to arise from German unification. These limitations make the sample rather short, but our econometric methods, in particular the instrumentalisation following Holtz-Eakin, Newey and Rosen (1988), are designed to cope with small time dimensions.

We transformed all data into log first differences to generate growth rates. Then, to account for the potential role of global shocks that may create uninsurable output variability, we formulated the data for each country relative to the global aggregate. In the setup of the panel, we multiplied the data of each country by its population weight. The description of variables from the PWT data base and the list of countries are given in the data appendix.

We then proceeded to the panel estimation of the vector autoregressions given in (4.9), using the method suggested by Holtz-Eakin, Newey and Rosen, as described in the previous subsection and the technical appendix. In the estimation, we included time-specific fixed effects to account for any remaining cross-sectional dependence and individual-specific fixed effects.

We used standard information criteria to determine the lag length of the VAR-model. Those generally suggested 1-2 lags. As an additional test of specification, we used the fact that, as k tends to infinity, $\beta(k)$ should converge to the unconditional β , i.e. the vector of coefficients of the simple panel regressions on Δgdp of $\Delta gdp - \Delta gnp$, $\Delta gnp - \Delta c$ and Δc respectively. This test generally required us to impose two lags which we used throughout.

We then inverted the VAR to generate forecast errors according to (4.6). We now discuss the results of this exercise.

4.4 Results

In the selection of countries we used for our investigation, we deliberately only included industrialised economies. This is to ensure that countries are sufficiently homogenous to warrant treatment in a single panel estimation. Our panel also includes several interesting sub-groupings and we will report results for these throughout. These sub-groups include the G7, the EU 15 and a core group of European economies. Again, the appendix provides more detail.

Table 4.1 provides the relative contributions of the intertemporal and the cross-sectional channels at various forecast horizons.

It is a first interesting feature of our results that the relative contribution of the channels does not vary over time. To save space, table 4.1 reports, the results for the one and three year horizons only but the findings at other horizons are virtually identical. This is a remarkable

result that we found to be extremely robust to changes in the model specification. It may seem surprising that short-term and long-term risks are equally well insured. Intuitively, one might expect that various forms of capital market imperfections would lead to a 'short-term bias' in consumption insurance. This does not seem to be the case.

Our results may also suggest that a fully specified dynamic econometric model such as the one put forward in this paper is required to get at the issue of dynamic risk sharing. Earlier contributions to the literature which admittedly were not primarily concerned with the dynamics of risk sharing, tended to use a short-cut to gauge risk sharing at different horizons: they typically look at data differenced at high and low frequencies. Using this method, Sørensen and Yosha (1998) find that the unsmoothed component at the three-years horizon is much larger (roughly 75 percent) than at the 1-year horizon (roughly 60 percent). In a similar way, Canova and Ravn (1996) report that lower frequency fluctuations in income are less insured than higher frequency fluctuations. The results in our paper, including those to be reported below for permanent and transitory shocks, demonstrate that these exercises provide a good estimate of how well-insured countries are against shocks of various degrees of persistence but not how well-insured they are against risks at different horizons.

The last column of table 4.1 reports the estimate of the unconditional model, i.e. practically a re-run of the Sørensen and Yosha (1998) procedure on our data. These unconditional estimates display the same pattern that was already found by Sørensen and Yosha (1998). Capital markets virtually do not matter for risk sharing, the bulk of insurance is provided through (intertemporal) self-insurance. Interestingly enough, our conditional estimates from the dynamic model find a somewhat more important role for international capital markets. However, once one takes account of the estimation uncertainty in the unconditional model, the respective results are not too far apart.

Overall, we find that the conditional estimates eventually converge to unconditional ones - at least after taking account of the relatively large estimation uncertainty in the unconditional model. This is reassuring as it provides a check of specification of our dynamic model as has been suggested in section two.

The VAR-based approach we have suggested in this paper allows us to examine an important assumption that underlies the ASY-approach: if the GDP-GNP differential and the GNP-Consumption differential actually serve as buffers for shocks to output, they should be driven by exactly the same shocks that drive GDP. In other words; the notion underlying ASY and the related literature is that shocks originate in output fluctuations and get smoothed at various levels. But the various aggregates, i.e. the GDP-GNP differential and the GNP-C differential that act as buffers, should not themselves be the source of shocks.

We can examine this assumption by conducting a principal components analysis of the shocks to our econometric model. If the presumption underlying the ASY approach is correct, then there should be a single dominant principal component in the reduced form errors that we get from the estimation of the VAR. Furthermore, this principal component should be highly correlated with innovations in the Δgdp -equation of our model but virtually uncorrelated with innovations in the other two equations.

In table 4.2, we give the share of the total variation in $[\Delta gdp, \Delta gdp - \Delta gnp, \Delta gnp - \Delta c]$ that is explained by the first principal component of Ω . As it turns out, we *do* find a dominant principal component in the reduced-form errors for all groupings of countries that we examine. We then also calculated the correlation of this principal component with unexpected innovations in the Δgdp -equation, i.e. ε_{1t} as well as the $\Delta gdp - \Delta gnp$ - and $\Delta gnp - \Delta c$ -equations, ε_{12t} and ε_{13t} respectively. These correlations are given in columns 2-4 of table 4.2. Our results suggest that, indeed, shocks to Δgdp drive the joint dynamics of $[\Delta gdp, \Delta gdp - \Delta gnp, \Delta gnp - \Delta c]$. This is a very important finding as it demonstrates the validity of our method and the static versions of it that have been used in ASY (1996), Sørensen and Yosha (1998) and Mélitz and Zumer (1999).

4.4.1 Permanent and transitory shocks

In table 4.3 we provide forecast error decompositions for the different sources of income uncertainty, i.e. permanent and transitory shocks.

These decompositions are similar to the unconditional dynamic results we reported in table 4.1 in that they do not vary over time. However, our results also reveal that there are important differences in the way that the various channels contribute to the sharing of risks that arise from different sources of shocks.

Firstly, permanent shocks are insured to a much lesser extent than transitory shocks. This finding is in line with earlier results in Canova and Ravn (1996) who also found that low-frequency risks seem to be insured less than high frequency fluctuations. In fact, when the panel VECM is estimated with all countries included, transitory shocks are found to be almost perfectly smoothed. We note that, very much as in the unconditional case, the forecast horizon does not matter for the extent of total insurance nor for the relative role of the channels.

Secondly, once we consider the channels by which these shocks get insured, we find that insurance against transitory fluctuations is almost exclusively achieved through the intertemporal channel, whereas, in line with the findings by Sørensen and Yosha (1998), the role of capital markets remains limited.

The results for transitory shocks tie in with recent empirical research by Kraay, Loayza, Serven and Ventura (2000) that suggests that the international component of most countries' portfolios is heavily biased towards loans and bonds. On the theoretical side, Baxter and Crucini (1995) have shown that the full risk sharing allocations that ensue as equilibria in models with complete markets can be approximated by models that feature only non-state-contingent assets. This result holds as long as shocks are not too persistent. Our results highlight the empirical relevance of the Baxter and Crucini study: even though individual countries' international portfolios show a huge home bias, they seem sufficiently diversified to achieve almost full insurance against transitory output risks. This insurance seems to be achieved largely with bonds and international credit rather than equity or other state contingent assets.

The results we obtained for permanent shocks are particularly interesting in two respects. First, when the model is estimated with data from all countries, the intertemporal and the cross-sectional channels play almost equal roles. Certainly, the bulk of permanent income risk remains uninsured, but the relative contributions to the amount of insurance that is eventually achieved is roughly equal for the cross-sectional and intertemporal channels. In particular, it is noteworthy that the intertemporal channel matters at all for the insurance against permanent risks. Models in which only the expected path of income matters for the savings decision will not be able to rationalize this feature of the data. Rather, income variability appears to matter in this case. Athanasoulis and Shiller (2000) have shown how the extent of observed precautionary saving depends on the degree of incompleteness of markets for claims on national income. Accordingly, we interpret our finding as evidence of precautionary savings.

When the model is estimated with only a subgroup of countries, our results are generally confirmed. One particular point may be worth mentioning, though:

The role of the cross-sectional channel, i.e. international capital markets, for the insurance against permanent shocks is less pronounced in all of the sub-groups than it is when the model is estimated with all 23 countries. The sub-groups are more homogenous in terms of country-size than is the whole panel. Our results could suggest that risk sharing through international equity markets is more pronounced between countries of different size. In this respect, our results are in line with Lane (2000) who has found that smaller countries tend to hold more foreign equity than do larger countries.

Summarizing this section, we can say that, in annual data, permanent output fluctuations account for just below eighty percent of total output variability and that only twenty percent of these fluctuations are insured. Hence, we find that at least 60 percent of total output

variability is uninsured. This is in line with the results obtained by Sørensen and Yosha. Our findings complement theirs in that it seems that most of this uninsured component is due to uninsured permanent shocks. This raises the question why international capital and credit markets do so poorly in insuring people's consumption against permanent shocks in income.

4.5 Conclusion

In this paper, we have investigated in which way industrial countries insure against output fluctuations. In so doing, we have offered two important novelties.

The first is that we consider how risks are insured at various horizons, thus providing a dynamic version of the method first proposed by Asdrubali, Sørensen and Yosha (1996). Our results corroborate the notion of a home bias in international risk sharing, even for forecast horizons as low as a year.

Secondly, we also find evidence that an important share of the variation in idiosyncratic output and consumption may be of a permanent nature. Permanent shocks need to be insured through perpetual claims. The French-Poterba observation of a home bias in international equity portfolios may suggest that most countries are badly insured against permanent fluctuations in their income streams. Once we allow for non-stationarities in our data set, our findings are consistent with this view: there is generally little insurance against permanent shocks but transitory risk in output is almost completely insured, mainly via national and international credit markets.

The second finding is in line with recent empirical evidence that suggests that international portfolios do not only display a home bias but are also severely biased towards non-state contingent assets such as bonds and loans. A theoretical rationalization for our results may be given by Baxter and Crucini (1995), who demonstrated that full risk-sharing allocations can be approximated quite well in economies with imperfect capital markets as long as shocks are not too persistent.

Our aim in this paper was to draw a map of an area of our ignorance, i.e. how countries share risks at various time horizons. We have not put forward any particular theory of what the intertemporal pattern of risk sharing should look like. However, any theoretical model of the home bias should also reproduce the fact that the relative importance of risk sharing channels does not vary over time. This may, for example, be an important restriction on transaction-cost based explanations of the home bias as the presence of (fixed) transactions costs may well imply that the relative roles of intertemporal smoothing and cross-sectional

insurance vary with the forecast horizon.

Another issue that our results may raise is why permanent and transitory shocks are insured in such different ways. Apparently, permanent shocks are much harder to insure internationally than are transitory shocks. Why this should be the case is not immediately clear but it is what the data tell us. One intuition is that there is that the risk of reneging - on the

part of the insurer - is much higher with permanent shocks. We plan to further address this question in future research.

4.6 Appendix

4.6.1 Data Appendix

All data are from the Penn World Tables release 5.6

Variable	PWT-code	Line#
GDP	cgdp	9
GNP	rgnp	27
C	cc	10
Population	pop	1

The sample range is 1970-90.

List of Countries:

1. Canada, 2. United States, 3. Japan, 4. Austria, 5. Belgium, 6. Denmark, 7. Finland, 8. France, 9. Germany (West), 10. Greece, 11. Iceland, 12. Ireland, 13. Italy, 14. Luxemburg, 15. Netherlands, 16. Norway, 17. Portugal, 18. Spain, 19. Sweden, 20. Switzerland, 21. United Kingdom, 22. Australia, 23. New Zealand

G7: countries #1,2,3,8,9,13,21

EU 15: countries #4,5,6,7,8,9,10,11,12,13,14,15,17,18,19,21

EU core: countries #4,5,8,9,14,15

4.6.2 Technical Appendix

Anderson and Hsiao (1982) were the first authors to present a solution to the problem of estimating dynamic panel data models. For the case of a univariate AR(1), they suggested first differencing the model to get rid of the country-specific effects f_i and then using $(\Delta X_{i,t-2} - \Delta X_{i,t-3})$ or simply $\Delta X_{i,t-2}$ as an instrument for $(\Delta X_{i,t-1} - \Delta X_{i,t-2})$. These instruments will not be correlated with $(u_{i,t} - u_{i,t-1})$. This instrumental variables (IV) estimation method leads to consistent but not necessarily efficient estimates of the parameters of the model. In the sequel, several other studies (e.g. Arellano and Bond, 1991) suggested instruments leading to more efficient estimates. The above-mentioned problems are not specific to VARs.

In a landmark paper, Holtz-Eakin, Newey and Rosen (1988) - HNR for short - explained how to estimate VARs in a panel framework and proposed an IV type estimation procedure which we will now briefly explain.⁵

The specification of (4.8) as a projection implies that the error term u_{it} satisfies the orthogonality condition

$$E[\Delta X'_{is} u_{it}] = E[f'_i u_{it}] \quad s < t \quad (4.13)$$

We can exploit these orthogonality conditions to identify the parameters of the model. Taking first differences on (4.8), we obtain

$$\Delta X_{it} - \Delta X_{it-1} = \Lambda_t + \sum_{l=1}^p \Phi_l (\Delta X_{i,t-l} - \Delta X_{i,t-l-1}) + v_{it} \quad (4.14)$$

$$i = 1, \dots, K; t = p+2, \dots, T \quad (4.15)$$

where

$$\Lambda_t = \lambda_t - \lambda_{t-1} \quad (4.16)$$

$$v_{it} = u_{it} - u_{i,t-1}$$

We will now discuss identification of the parameters of the transformed equation (4.14) and then see how the original parameters can be recovered.

The orthogonality conditions of equation (4.13) imply that the error term of the transformed equation (4.14) satisfies the orthogonality condition

⁵Holtz-Eakin et al. deal with the more general case of an interacted country-specific and time effect $\lambda_i f_i$ and with time-varying coefficients.

$$E[\Delta X'_{it} v_{it}] = E[f'_i v_{it}] \quad s < t - 1 \quad (4.17)$$

Therefore,

$$Z_{it} = [e, (\Delta X_{i,t-2} - \Delta X_{i,t-3}), (\Delta X_{i,t-3} - \Delta X_{i,t-4}), \dots, (\Delta X_{i2} - \Delta X_{i1})]$$

qualify as instrumental variables. The original parameters are identified if $T \geq p + 3$.⁶ Note that the number of instruments increases with t . Thus, the HNR estimator is more efficient than an IV estimator based on once-lagged endogenous variables alone (as in Anderson and Hsiao, 1982).

Estimation yields the coefficients $[\Phi_1, \dots, \Phi_p]$, and we can calculate the variance-covariance matrix Ω^* of the transformed system. Using (4.16), we are able to recover the variance-covariance matrix of the original system. The estimated coefficients $[\Phi_1, \dots, \Phi_p]$ can be used to obtain the coefficient matrices C_t of the moving average representation. Finally, we can compute the mean squared prediction error using (4.6) from which the results in the main text follow immediately.

⁶Alternatively, following Arellano (1989) we used "level" values $Z_{it} = [e, \Delta X_{i,t-2}, \Delta X_{i,t-3}, \dots, \Delta X_{i1}]$ as instruments in which case we also gain one more period for estimation because in this case identification only requires $T \geq p + 2$. The results, however are very similar which we consider a robustness check of our empirical strategy.

Table 4.1: Risk Sharing at various horizons

Country group		Forecast Horizon in years				Unconditional Model	
		1		3			
a) All	β_C	0.09	(0.004)	0.09	(0.004)	0.001	(0.03)
	β_S	0.29	(0.002)	0.29	(0.002)	0.32	(0.14)
b) G7	β_C	-0.03	(0.01)	-0.03	(0.01)	0.01	(0.03)
	β_S	0.51	(0.04)	0.51	(0.04)	0.33	(0.13)
c) EU15	β_C	0.004	(0.01)	0.004	(0.01)	-0.02	(0.04)
	β_S	0.19	(0.05)	0.19	(0.05)	0.21	(0.18)
d) EU core	β_C	-0.02	(0.02)	-0.02	(0.02)	-0.02	(0.04)
	β_S	0.33	(0.09)	0.33	(0.09)	0.34	(0.18)

Values in parentheses are standard deviations.

For the unconditional model these are asymptotic whereas for the dynamic model they have been obtained using 100 bootstrap replications.

Table 4.2: Share of first principal component and correlation with GDP shocks

	Variance explained by 1. PC	Correlation of 1. PC with		
		ε_{i1t}	ε_{i2t}	ε_{i3t}
All	91%	0.99	-0.09	0.00
G7	94%	0.99	-0.13	0.00
Euro 15	77%	0.99	-0.13	0.00
EU core	78%	0.97	-0.23	0.00

Table 4.3: Smoothing of permanent and transitory shocks

Country group		Type of shock		Variance share	
		permanent	transitory	of perm. shocks	
a) All	β_C	0.10 (0.005)	0.03 (0.01)		
	β_S	0.11 (0.03)	0.91 (0.40)	0.77 (0.03)	
b) G7	β_C	-0.05 (0.02)	-0.03 (0.03)		
	β_S	0.36 (0.06)	0.96 (0.32)	0.72 (0.05)	
c) EU15	β_C	-0.007 (0.02)	0.03 (0.03)		
	β_S	0.51 (0.07)	-0.62 (0.38)	0.70 (0.04)	
d) EU core	β_C	-0.15 (0.05)	0.11 (0.05)		
	β_S	0.04 (0.19)	0.62 (0.26)	0.49 (0.09)	

Standard errors (in parentheses) were obtained from
100 bootstrap replications of the model

Bibliography

- [1] Anderson, T. W., and Cheng Hsiao (1982) 'Formulation and Estimation of Dynamic Models Using Panel Data' *Journal of Econometrics*, vol. 18(1), Jan. 1982, pp. 47-82.
- [2] Artis, Michael J., and Mathias Hoffmann (2001) 'International Cake Sharing and Consumption Correlations', mimeo, European University Institute and University of Southampton.
- [3] Asdrubali, P., Sørensen, P. E. and Yosha, O. (1996). 'Channels of interstate risk sharing: United States 1963-90.' *Quarterly Journal of Economics*, vol. 111, pp.1081-1110.
- [4] Athanasoulis, S. and Shiller, R. J. (2000). 'World Income Components: Measuring and Exploiting Risk Sharing Opportunities.' *American Economic Review*, forthcoming.
- [5] Athanasoulis, S. and van Wincoop, E. (2000). 'Growth Uncertainty and Risk Sharing.' *Journal of Monetary Economics*, vol. 45, pp. 447-505.
- [6] Arellano, M. and Bond, S. (1991). 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations.' *Review of Economic Studies*, vol. 58, pp. 277-97.
- [7] Backus, D., Kehoe, P. and Kydland, F. (1992), 'International Real Business Cycles', *Journal of Political Economy* 100(4): 745-775.
- [8] Baltagi, B. H. (1995). *Econometric Analysis of Panel Data*. John Wiley and Sons.
- [9] Baxter, M. and Crucini, M. J. (1995), 'Business Cycles and the Asset Structure of Foreign Trade', *International Economic Review* 36(4):821-853.
- [10] Caballero, R. J. (1991). 'Earnings Uncertainty and Aggregate Wealth Accumulation.' *American Economic Review*, vol. 81, No. 4. (Sep., 1991), pp. 859-871.

- [11] Canova, F. and Ravn, M. (1996). 'International consumption risk sharing.' *International Economic Review*, vol 37, no. 3, pp. 573-601.
- [12] Chamie, N., DeSerres, A. and LaLonde, R. (1994). 'Optimum Currency Areas and Shock Asymmetry - a comparison of Europe and the United States.' Working Paper 94-1. Bank of Canada.
- [13] French, K. and Poterba, J. (1991). 'Investor diversification and international equity markets.' *American Economic Review*, vol. 81(2), pp. 222-26.
- [14] Holtz-Eakin, D., Newey, W. and Rosen, H. S. (1988). 'Estimating vector autoregressions with panel data.' *Econometrica*, vol. 56, no. 6 (November), pp. 1371-95.
- [15] Johansen, S. (1995). *Likelihood Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press. Oxford.
- [16] King, Robert, Plosser, Charles I., Stock, James H. and Watson, Mark W., (1991). 'Stochastic Trends and Economic Fluctuations.' *American Economic Review*, vol 81, pp. 819-40.
- [17] Kraay, A., Loayza, N., Servén, L., and Ventura, J. (2000), 'Country Portfolios' NBER working paper #7795.
- [18] Kraay, A., and J. Ventura (2000) 'Current Accounts in Debtor and Creditor Countries' *Quarterly Journal of Economics*, 1137-1166.
- [19] Lane, P. D. (2000). 'International investment positions: a cross-sectional analysis.' *Journal of International Money and Finance*, vol. 19., pp. 513-534.
- [20] Lewis, Karen K. (1999) 'Trying to Explain Home Bias in Equities and Consumption', *Journal of Economic Literature*, vol. 37(2), June 1999, pp. 571-608.
- [21] Mark, Nelson and Sul, D. (2001). 'Nominal exchange rates and monetary fundamentals: Evidence from a small post-Bretton Woods panel.' *Journal of International Economics*. vol. 53, pp. 29-52.
- [22] Mélitz, Jacques and Frédéric Zumer (1999), Interregional and International Risk-Sharing and Lessons for EMU, Carnegie-Rochester Conference Series on Public Policy 51, December 1999, pages 149-188.

- [23] Mundell, R. (1961), 'A Theory of Optimum Currency Areas', *American Economic Review*, 51, pp- 657-65.
- [24] Neusser, K. (1991). 'Testing the Long-Run Implications of the Neoclassical Growth Model.', *Journal of Monetary Economics*, vol. 27(1), February 1991, pp. 3-37.
- [25] Sørensen, B. E. and Yosha, O. (1998). 'International risk sharing and European Monetary Unification.' *Journal of International Economics*, vol. 45, pp. 211-38.

Appendix A

The GSOEP data

The GSOEP is a representative longitudinal sample of the population residing in Germany and contains socioeconomic information on private households.¹ It began in 1984 with a sample of 12,245 respondents in 5,921 households in the western states of Germany. It consisted of two randomly sampled subgroups. The German Subsample consists of people in private households where the head of household is not Turkish, Greek, Yugoslavian, Spanish, or Italian nationality (the five largest groups of foreign nationals). The Foreign Subsample consists of people in private households where the head of household is of Turkish, Greek, Yugoslavian, Spanish, or Italian nationality.² In 1990, already before official unification, the first wave of the East German Subsample was added. It includes individuals in private households where the head of household is/was a citizen of the German Democratic Republic. In 1995, a special sample of immigrants was for the first time interviewed. In 1998, a refreshment sample was independently drawn, covering the same target population as the existing four subsamples, thereby considerably increasing the overall sample size. By now there are thus five different subsamples which can be aggregated using design weights.

All household members aged 16 years or older are interviewed face-to-face and asked their personal situation. The household head³ additionally provides information on housing, housing costs, different sources of income and on children under 16 years old.

In this thesis, I only use the West-German and foreign subsamples for two main reasons.

¹ This section draws on several official publications by the German Institute for Economic Research (DIW) and on Wagner, Burkhauser and Behringer (1993).

² Note that the foreign sample consists of people who came to Germany in the 1950's and 1960's already and have therefore already assimilated to the native German population. In contrast, the immigrant sample (see below) includes foreigners who only recently came to Germany.

³ The head of the household is defined as the person who knows best about the general conditions under which the household acts and is supposed to answer this questionnaire in each given year.

The longitudinal dimension of these subsamples is longer than for the East German and immigrant samples which is crucial for the chapter on displacement. It is also crucial for the chapter on returns to education that contains information about place of childhood only in the 1985 wave and information on parental background only in the 1986 wave. Problems in comparability of the West- and East-German school systems also makes the West-German and foreign subsample a natural choice in the chapter on university dropouts.

